GHUM & GHUML: Generative 3D Human Shape and Articulated Pose Models

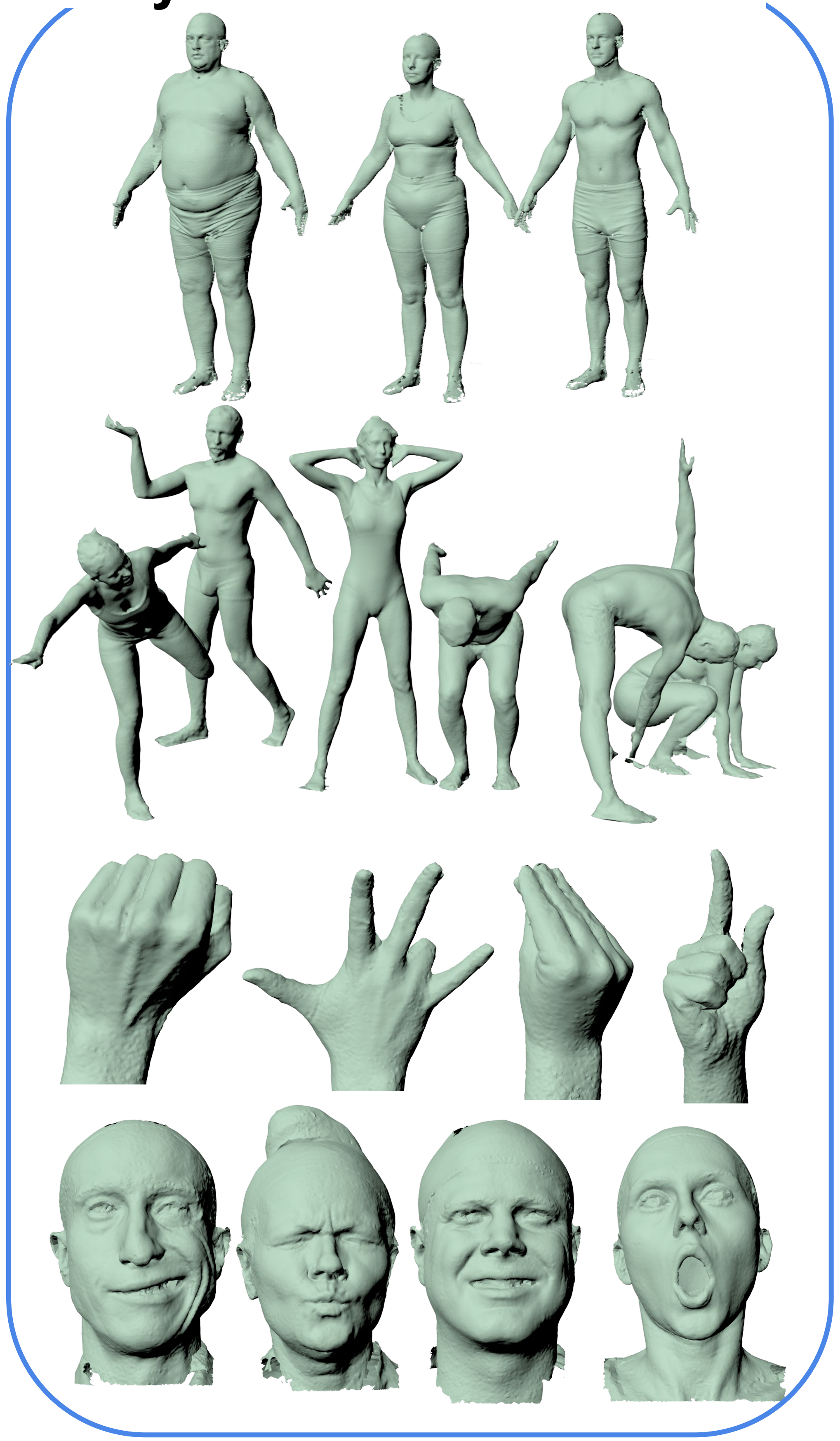
Hongyi Xu, Eduard Gabriel Bazavan, Andrei Zanfir,

William T. Freeman, Rahul Sukthankar, Cristian Sminchisescu

Google Research

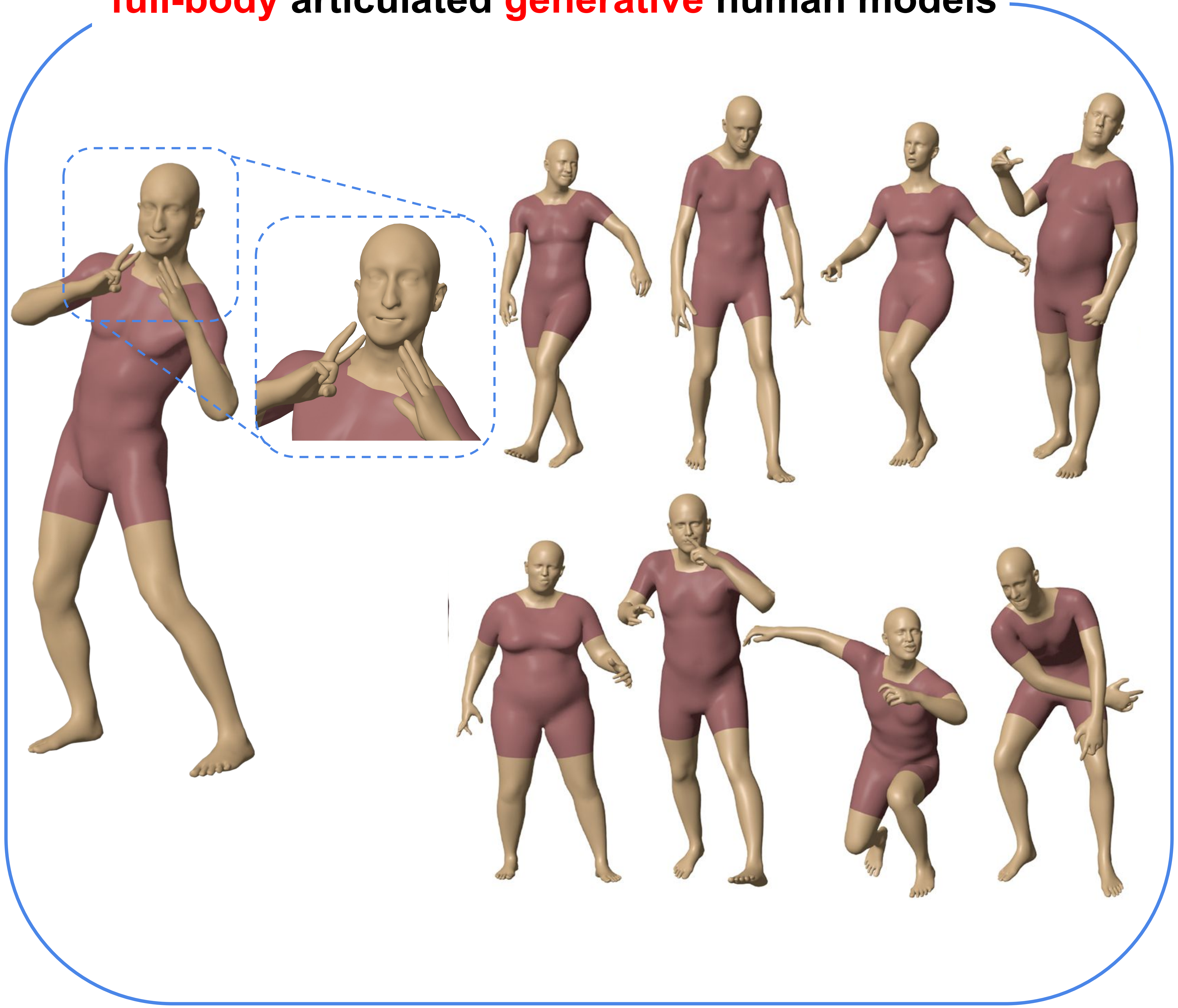
GENERATIVE HUMAN MODELING PIPELINE

dynamic human scans

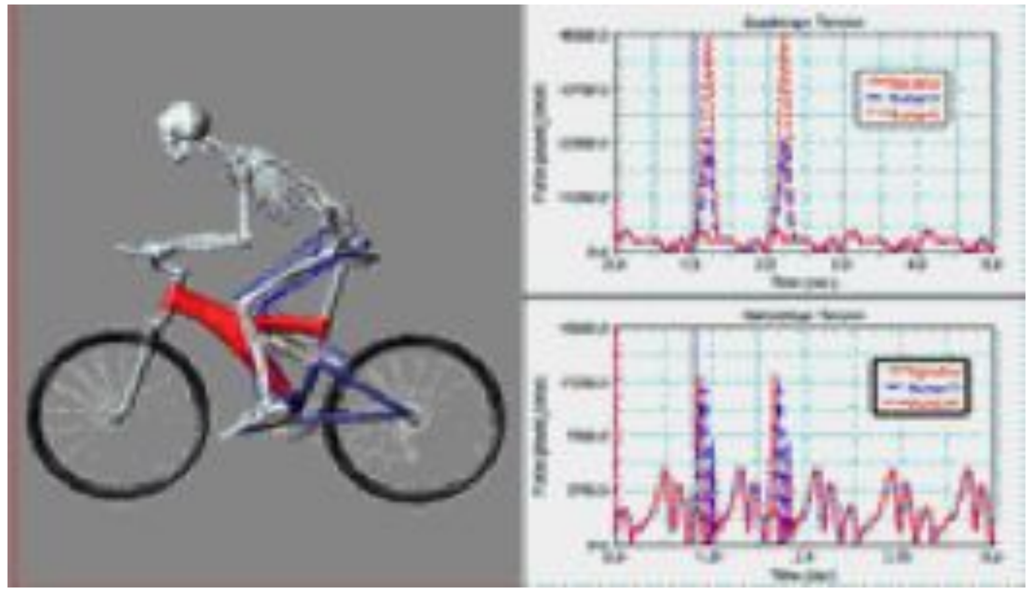


end-to-end deep learning

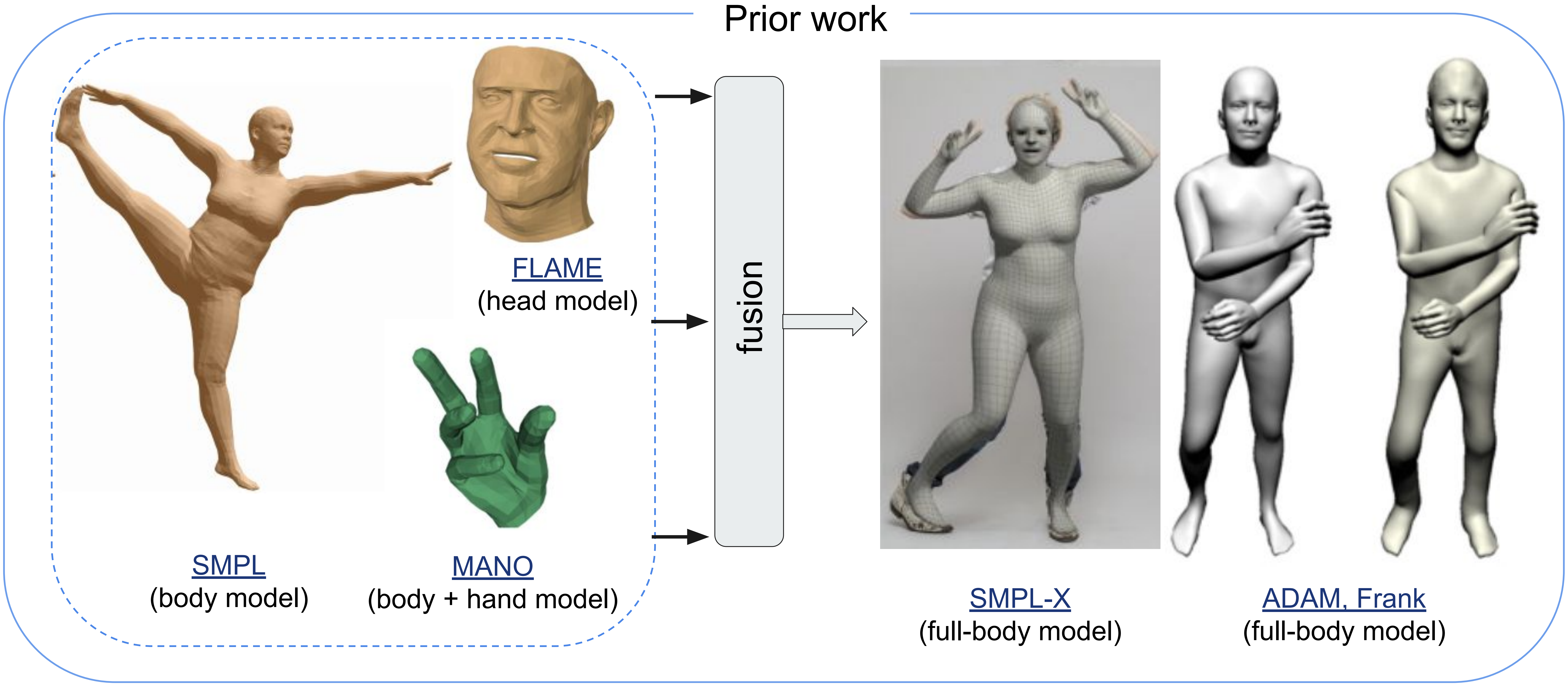
full-body articulated generative human models



Motivation: 3D Human Sensing



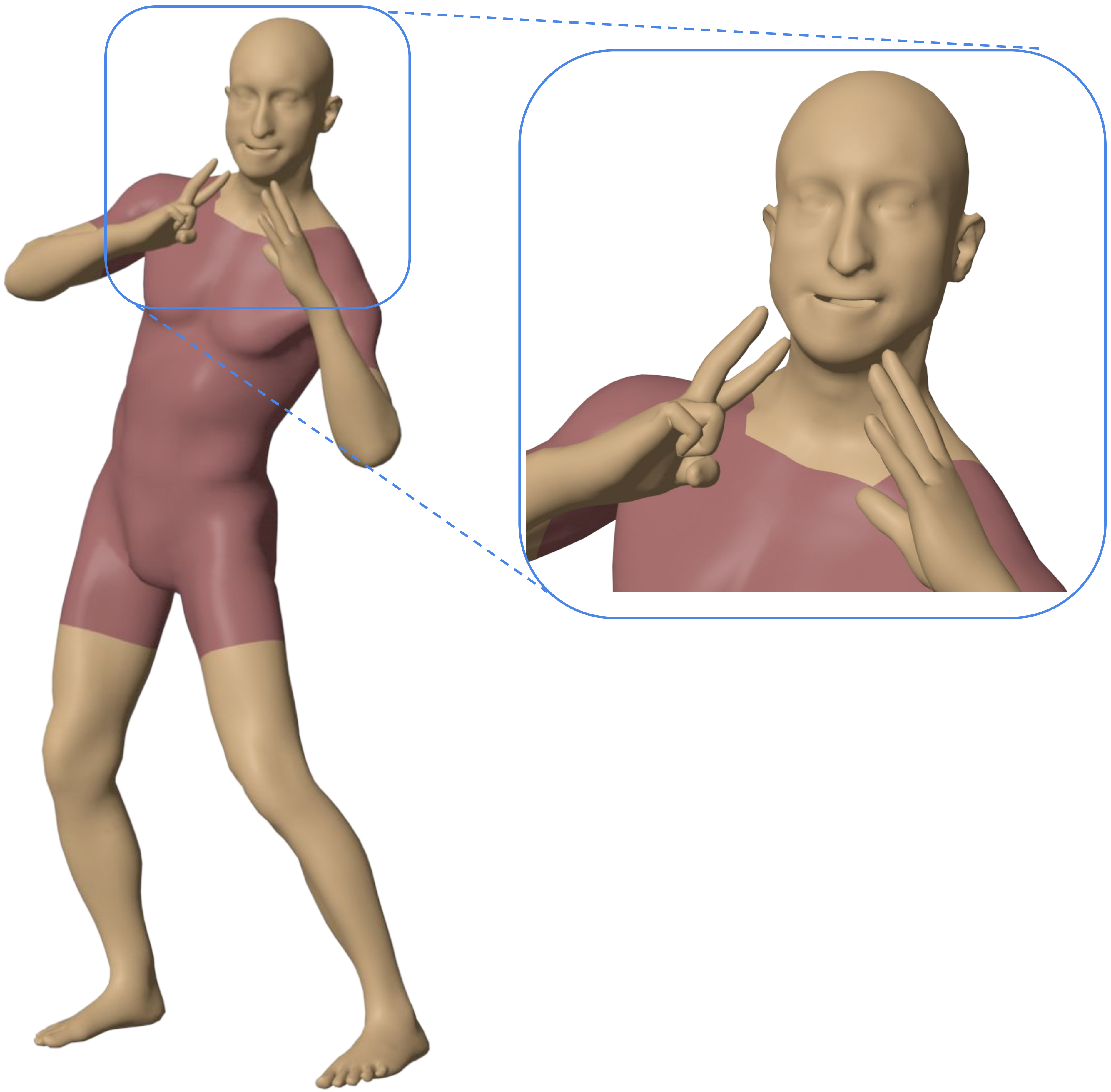
GENERATIVE HUMAN MODELS: PRIOR WORK



TOWARDS FULL-BODY GENERATIVE HUMAN MODEL

★ Full-body generative human model

★ All model components trained in a single consistent learning loop



END-TO-END TRAINING PIPELINE

Trainable modules:

- body shape VAE
- expression VAE
- pose-space deformation
- joint centers predictor
- skinning weights

Model parameters:

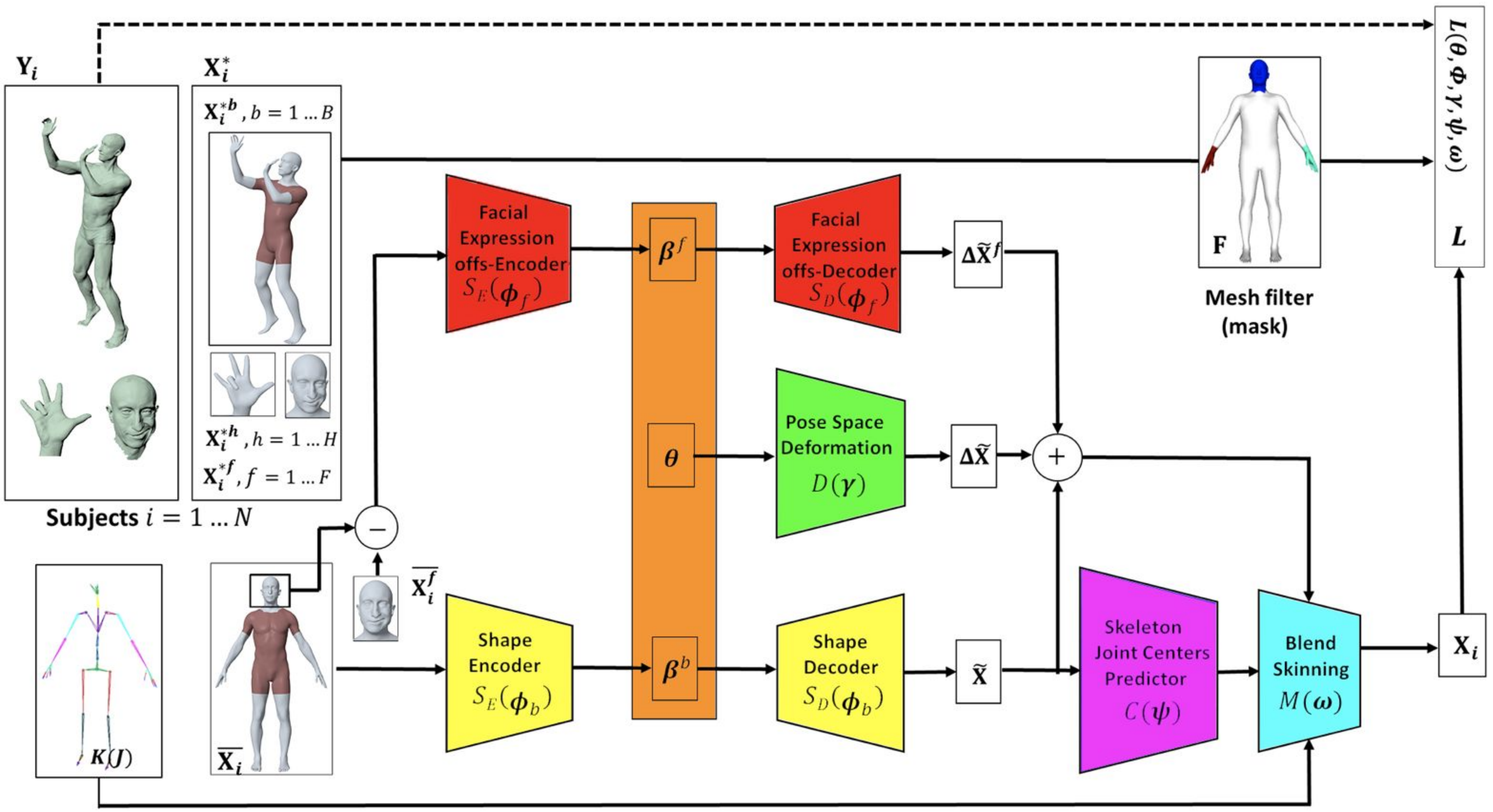
- latent codes: β^b, β^f
  pose: θ

Losses:

masked reconstruction loss

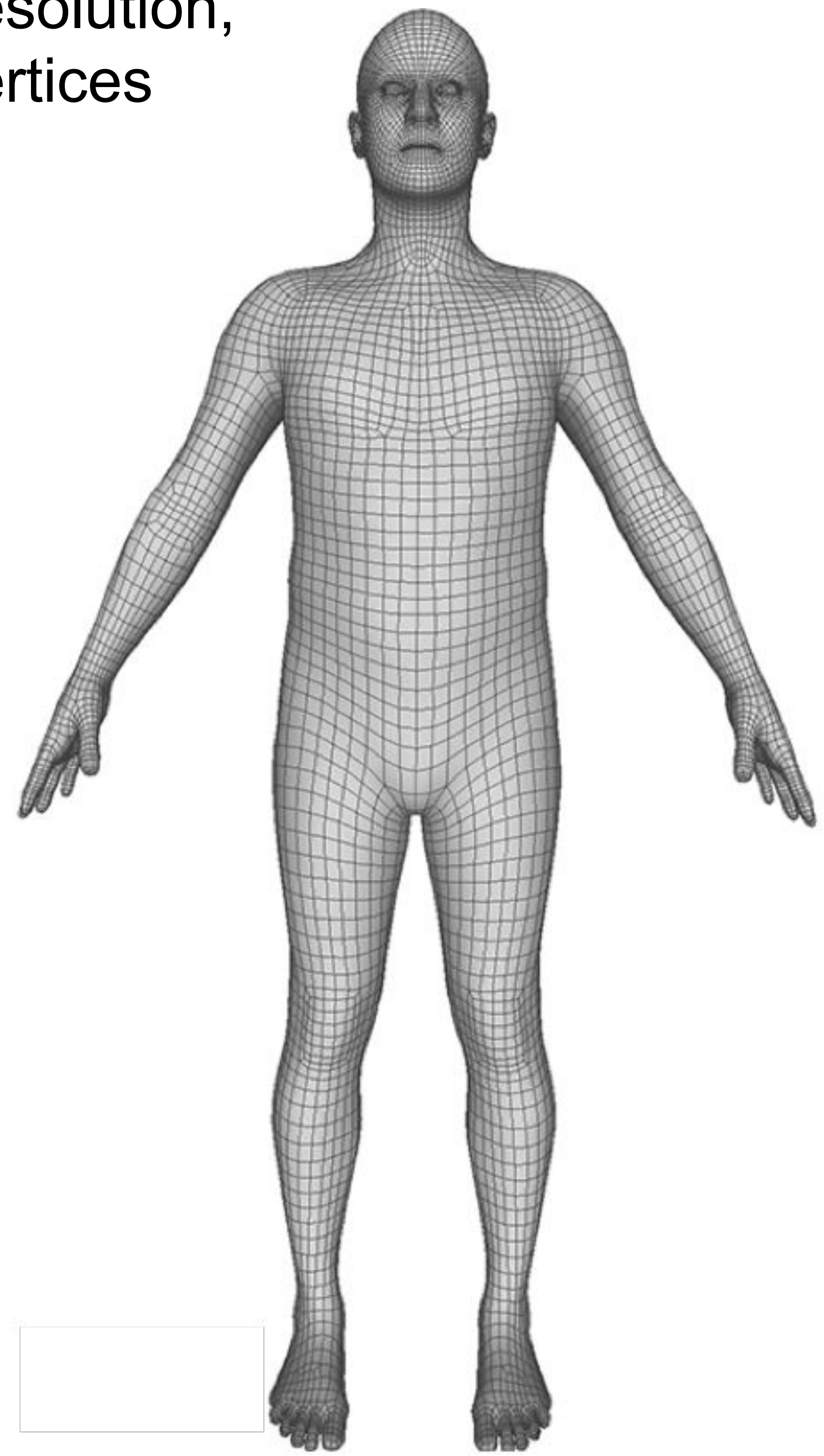
$$\frac{1}{N}\sum_{i}^{N} \|F_i(X_i^* - X_i(\beta^b, \beta^f, \theta))\|$$

+ parameters regularizers

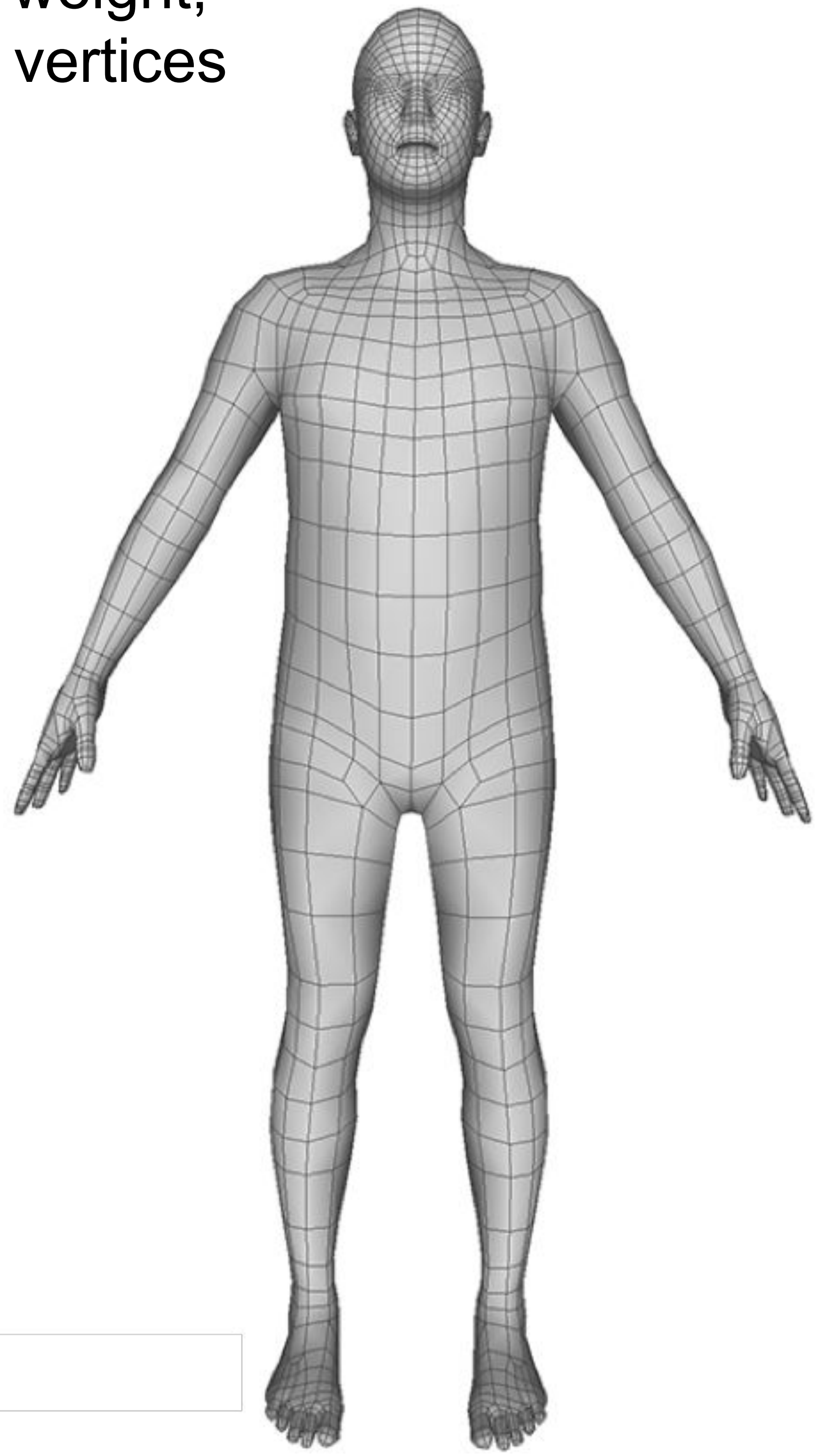


GHUM AND GHUML(ite): TEMPLATES

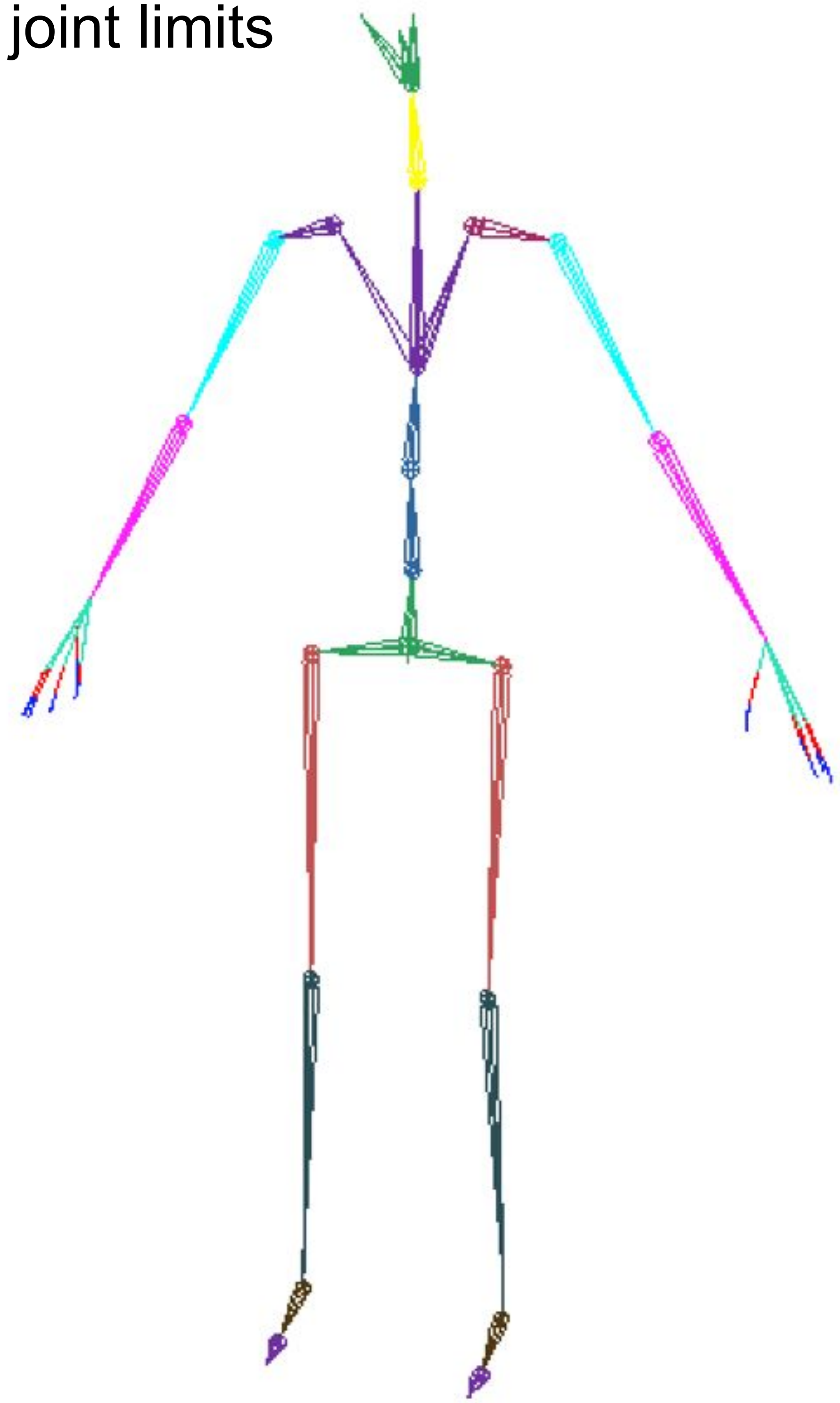
GHUM:
moderate-resolution,
10168 vertices



GHUML(-ite):
light-weight,
3190 vertices

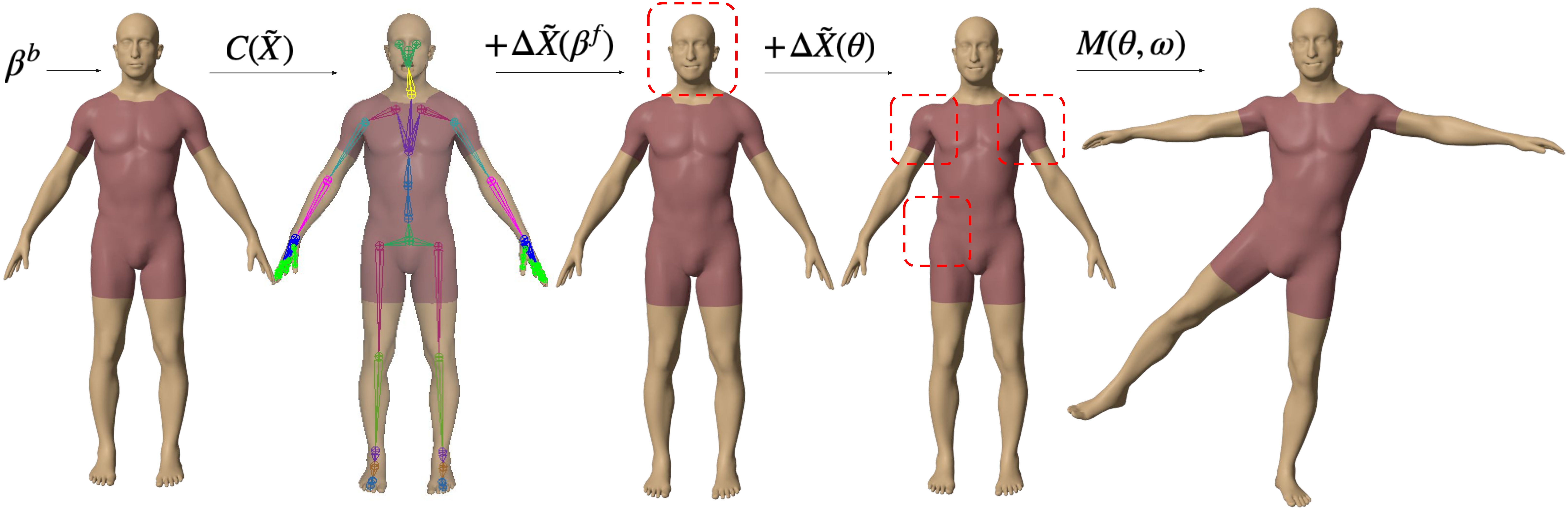


Shared skeleton:
minimal parameterized,
anatomical joint limits



GHUM AND GHUML(ite): MODEL FORMULATION

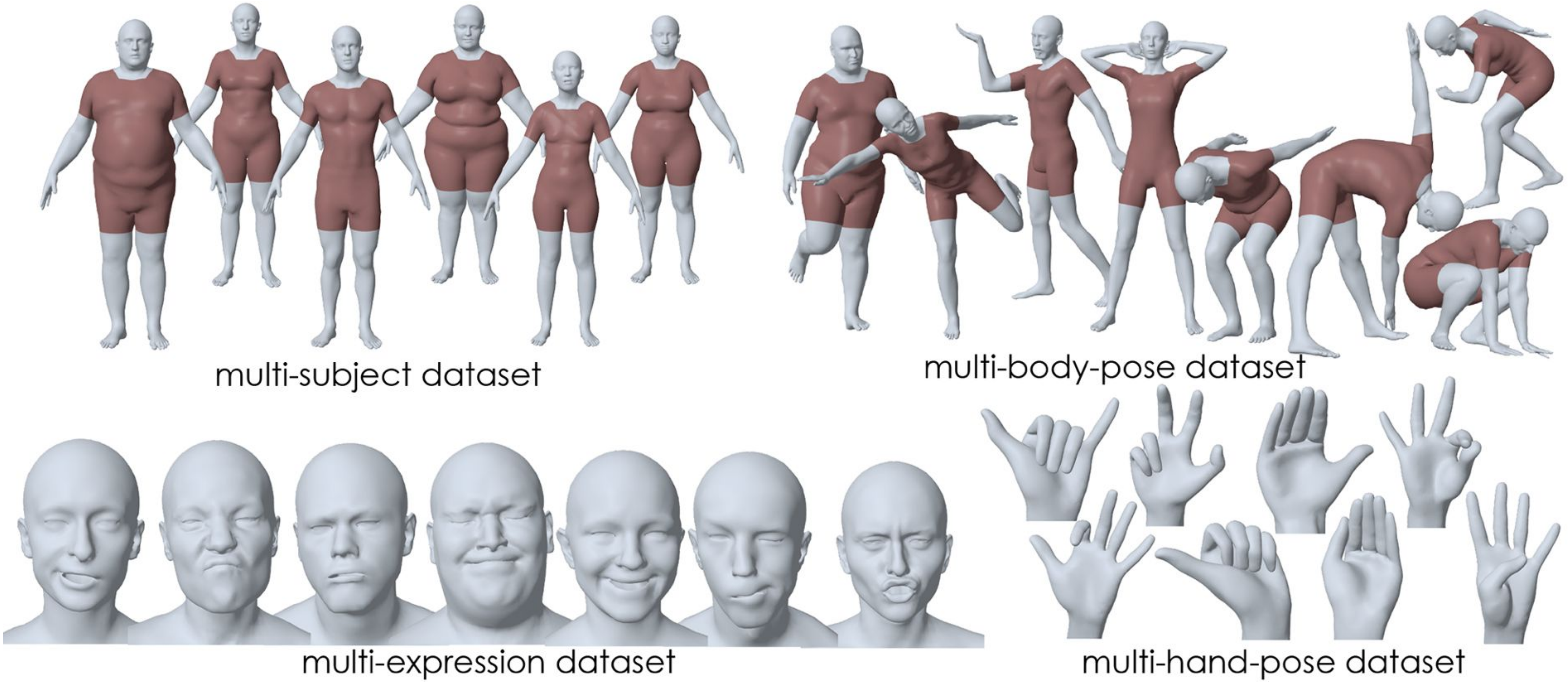
$$X(\beta^b, \beta^f, \theta) = M(\tilde{X}(\beta^b), \Delta\tilde{X}(\beta^f), \theta, \Delta\tilde{X}(\theta), C(\tilde{X}), \omega)$$



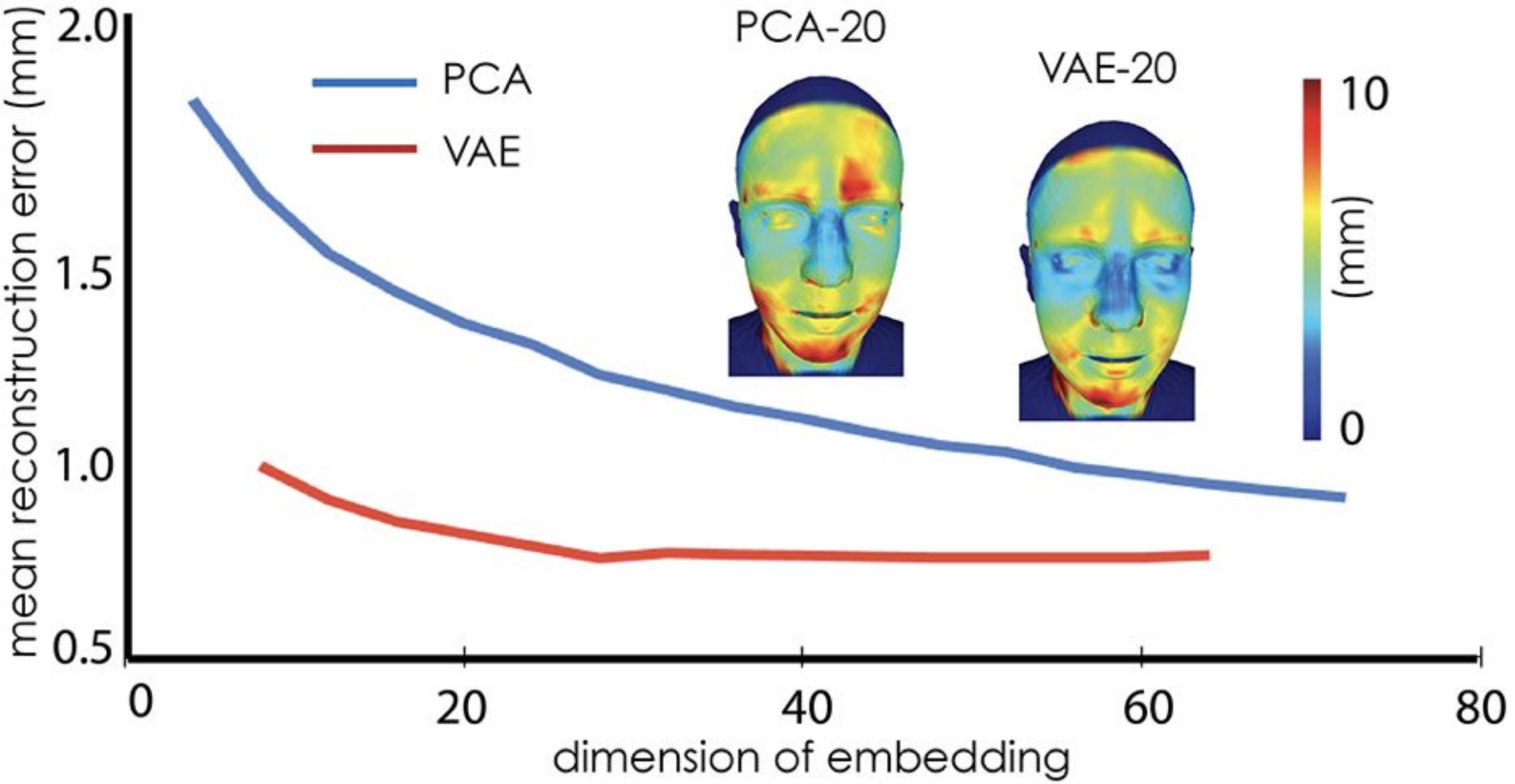
rest shape	joint centers	+ facial expression	+ pose-space deformation	output posed shape
$\tilde{X}(\beta^b)$	$C(\tilde{X})$	$\tilde{X}(\beta^b) + \Delta\tilde{X}(\beta^f)$	$\tilde{X}(\beta^b) + \Delta\tilde{X}(\beta^f) + \Delta\tilde{X}(\theta)$	X

REGISTRATION TO GHS3D + CAESAR

4.3K subjects, **60K** high-resolution dynamic human scans, including close-up head and hand scans.

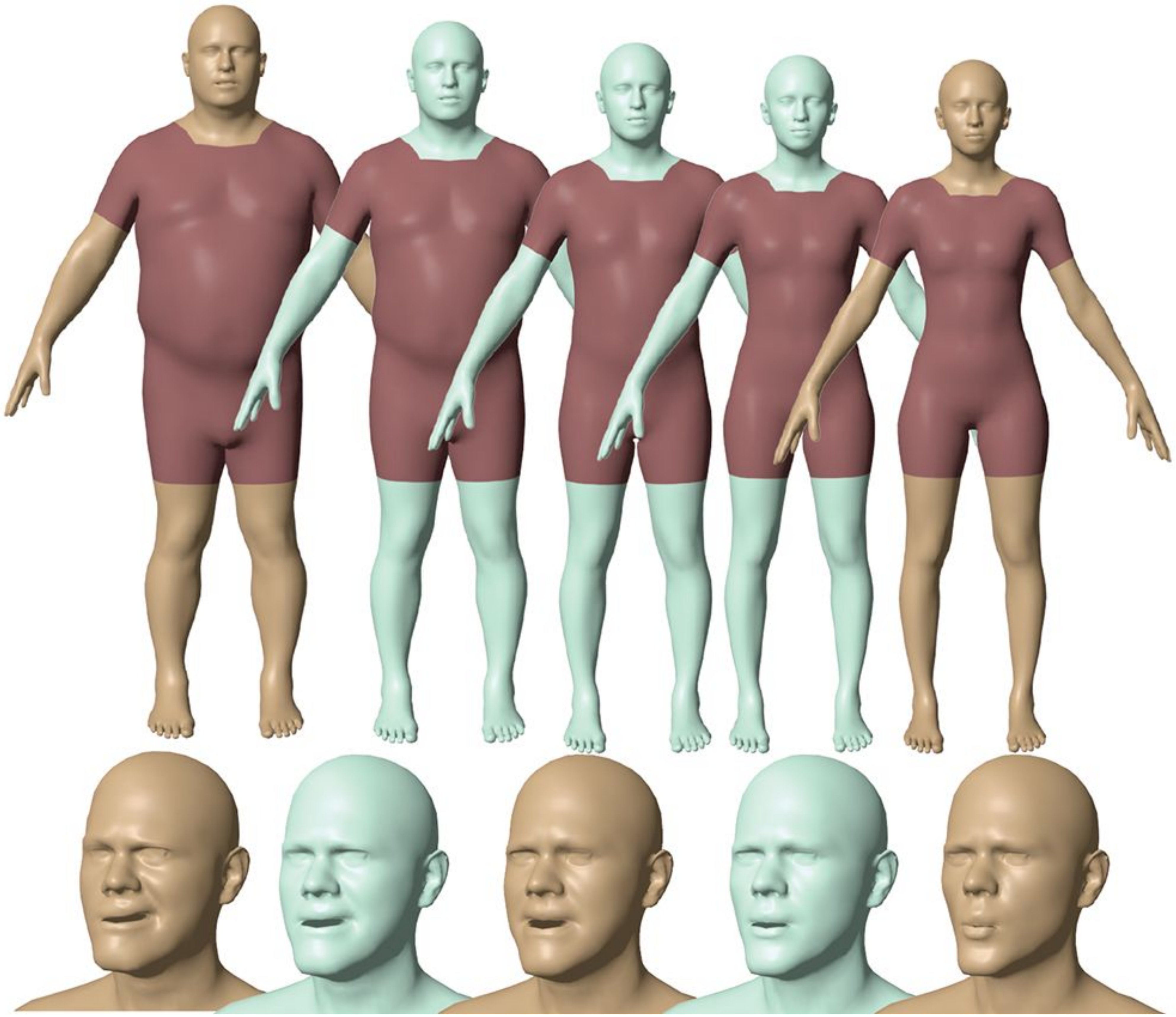


VARIATIONAL SHAPE AND EXPRESSION AUTOENCODER



Google Research

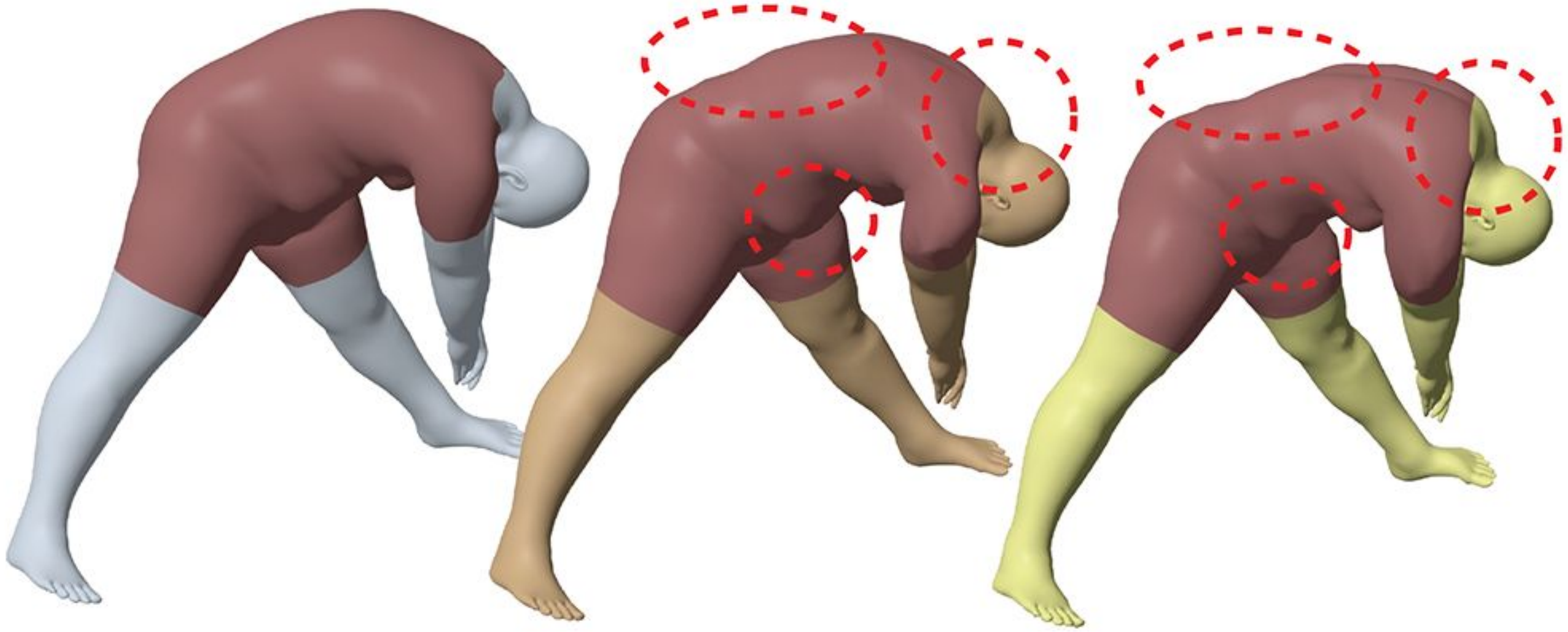
LATENT SPACE INTERPOLATION



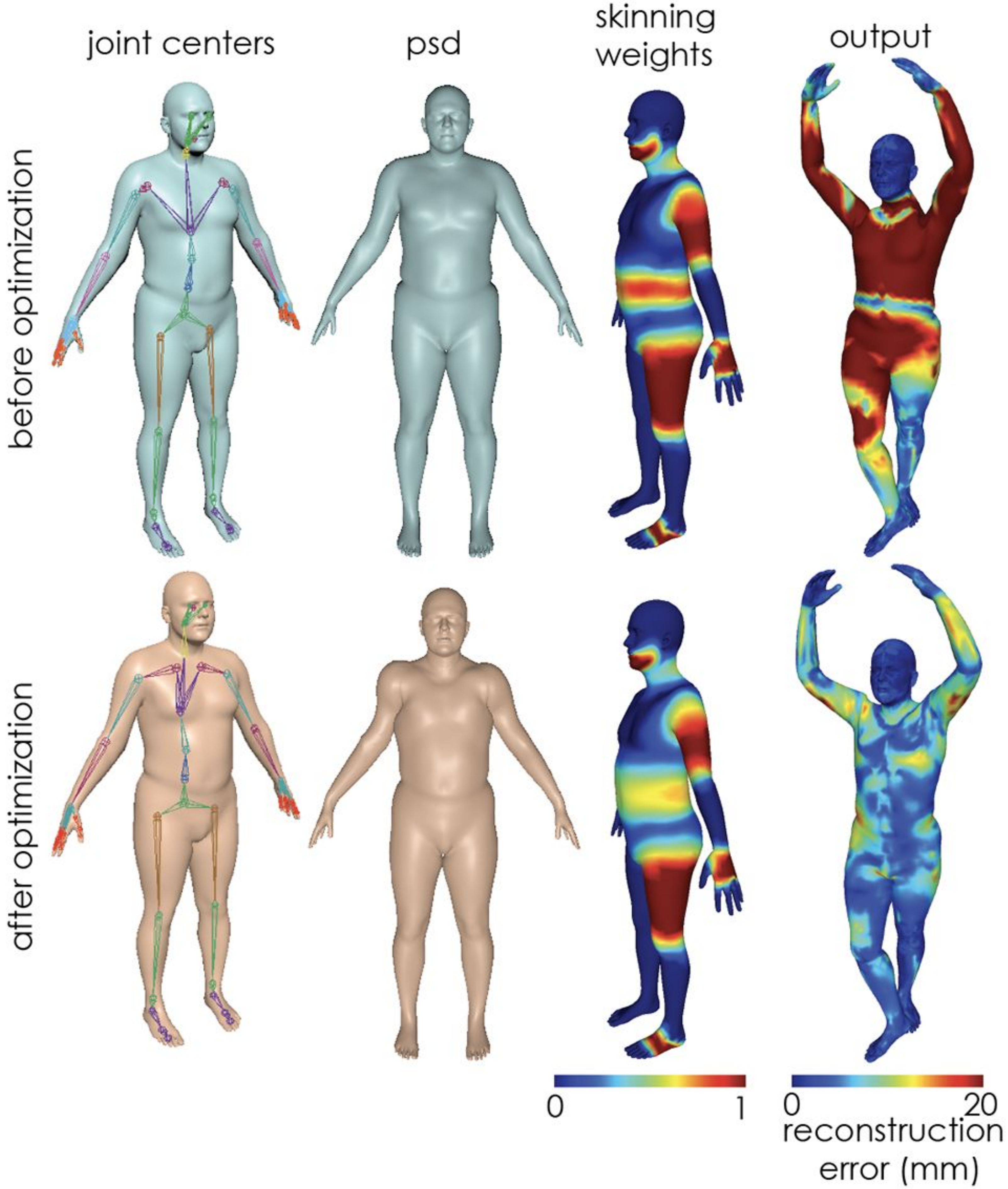
(brown: samples; green: interpolation)

Google Research

Optimized Skinning



registration GHUM with PSD GHUM without PSD



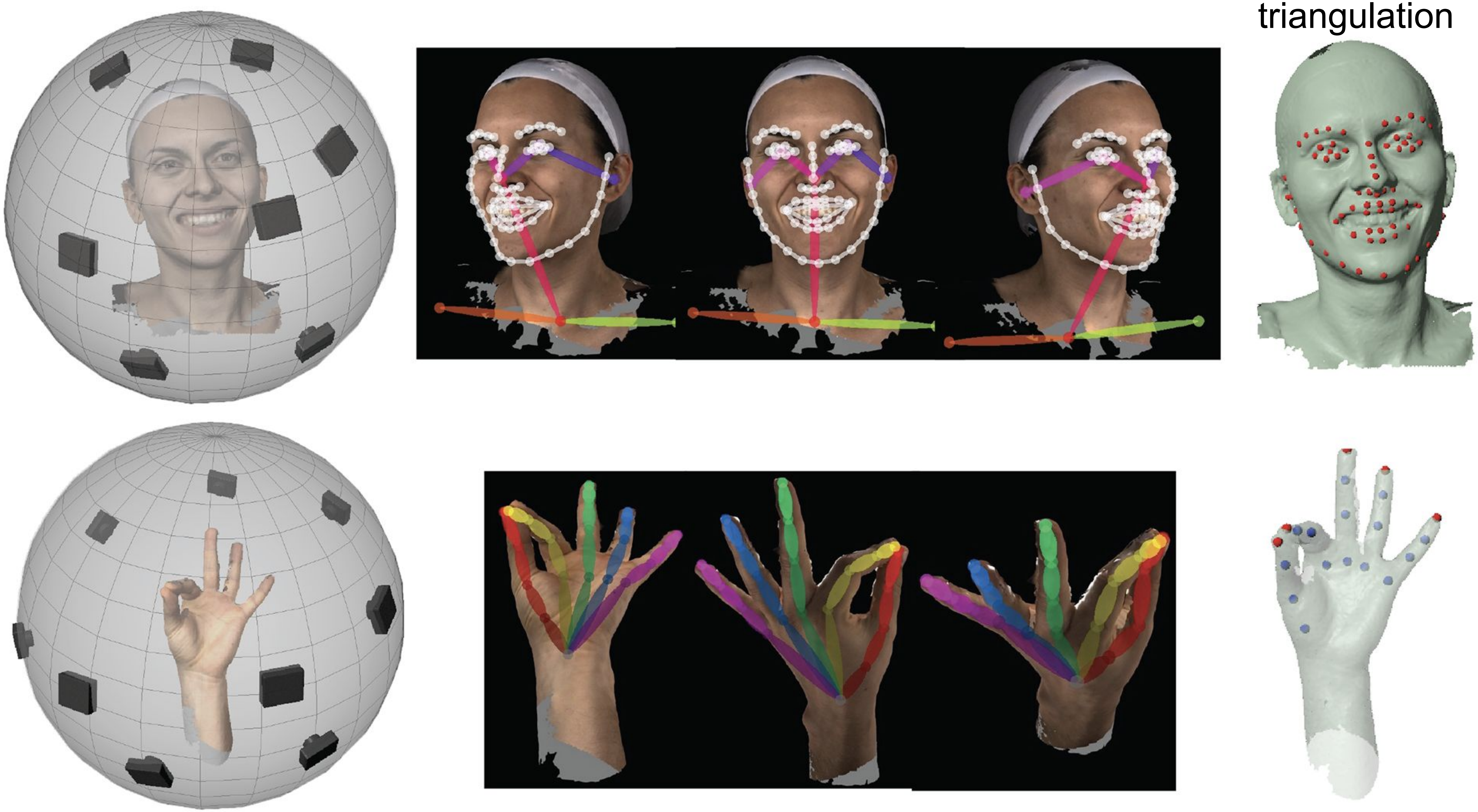
Google Research

AUTOMATIC 3D LANDMARKS DETECTION

Mutiview renderings

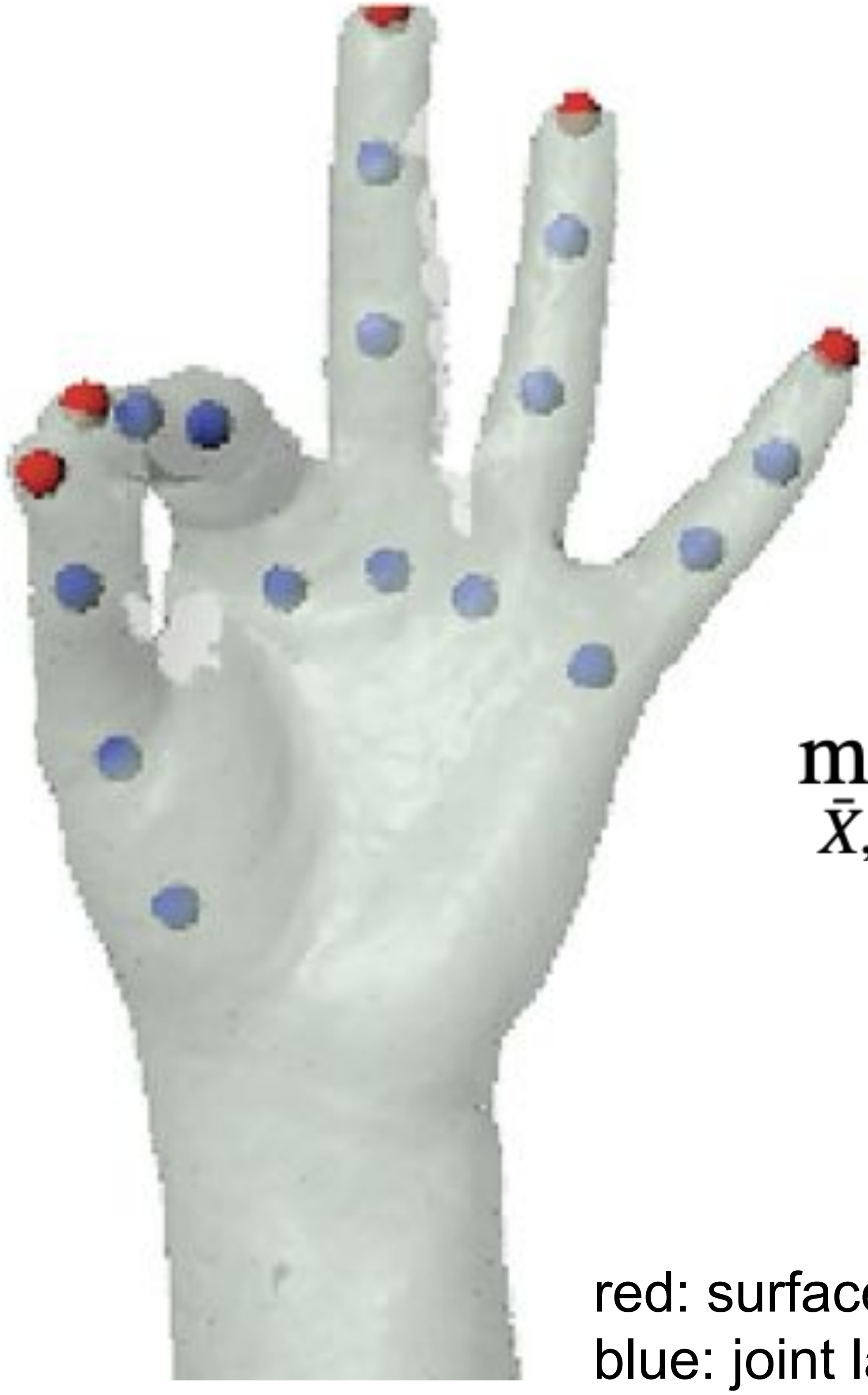
2D landmarks detection

3D landmarks triangulation

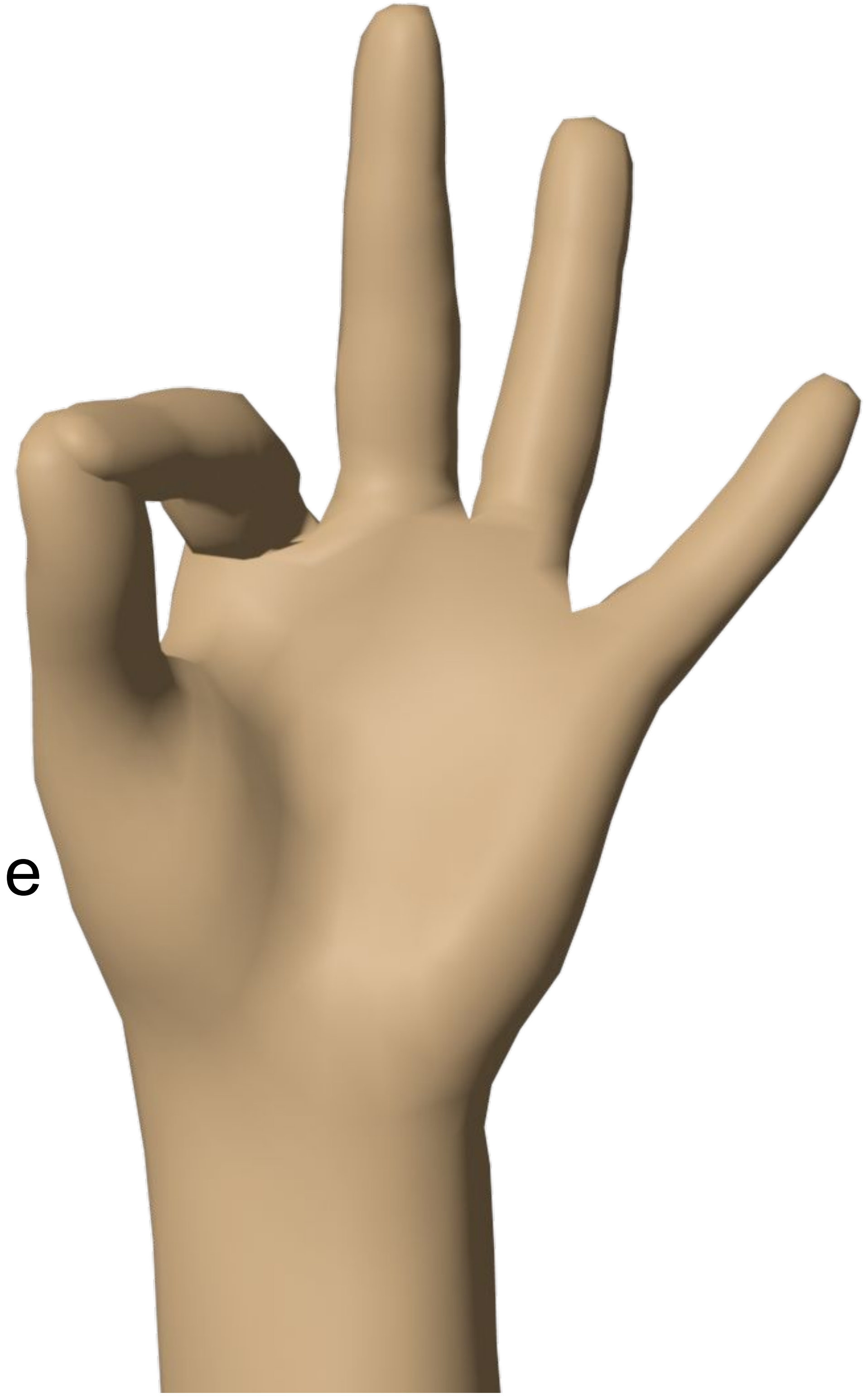


Google Research

ARTICULATED AS-CONFORMAL-AS-POSSIBLE REGISTRATION



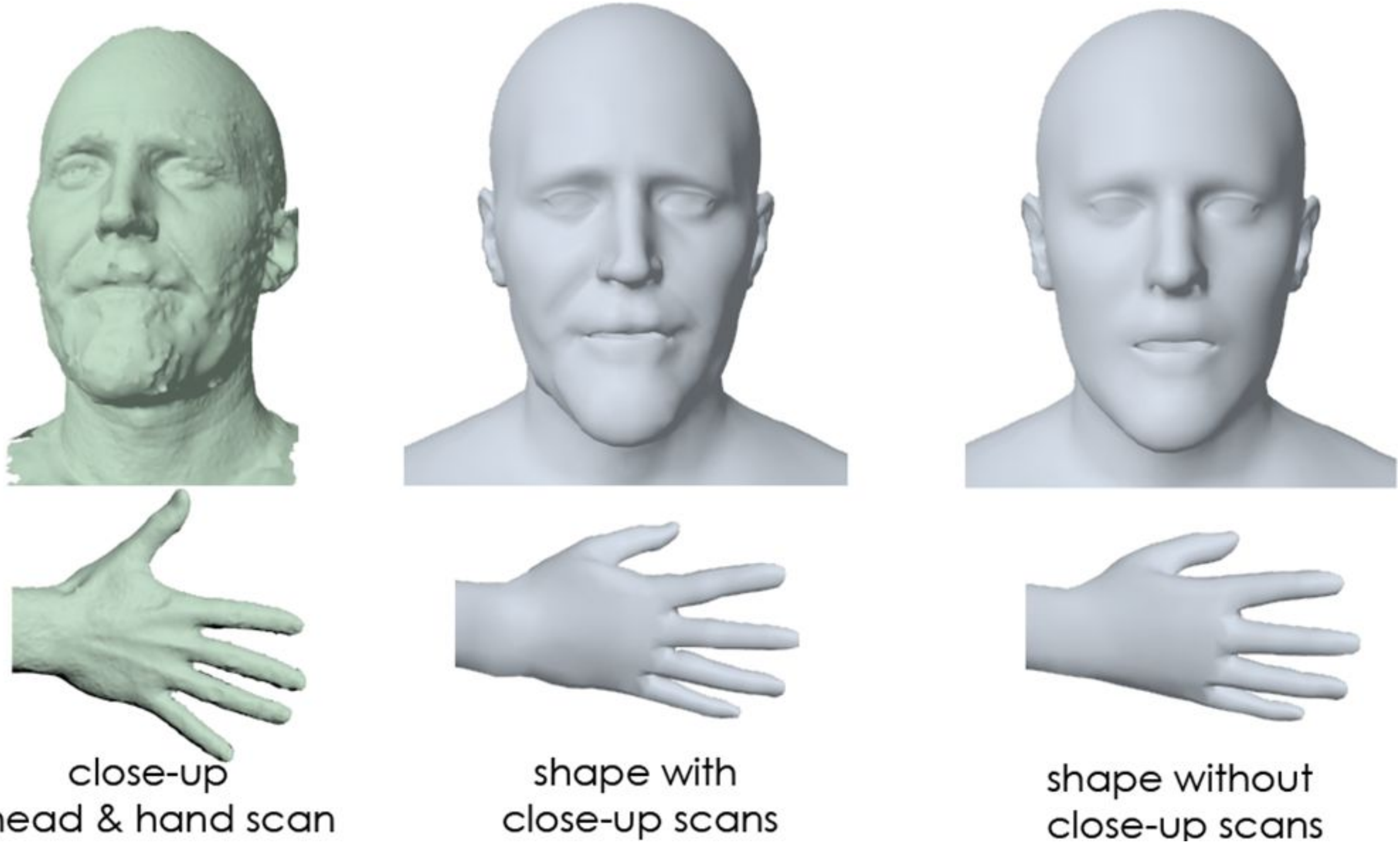
$$\min_{\bar{X},\theta} \xrightarrow[\text{+ shape prior: as-conformal-as-possible}]{\text{data term: ICP + landmarks}}$$



red: surface landmarks
blue: joint landmarks

FULL-BODY SHAPE ESTIMATION

We estimate full-body shape at A pose by fusing body scan with close-up head and hand scans

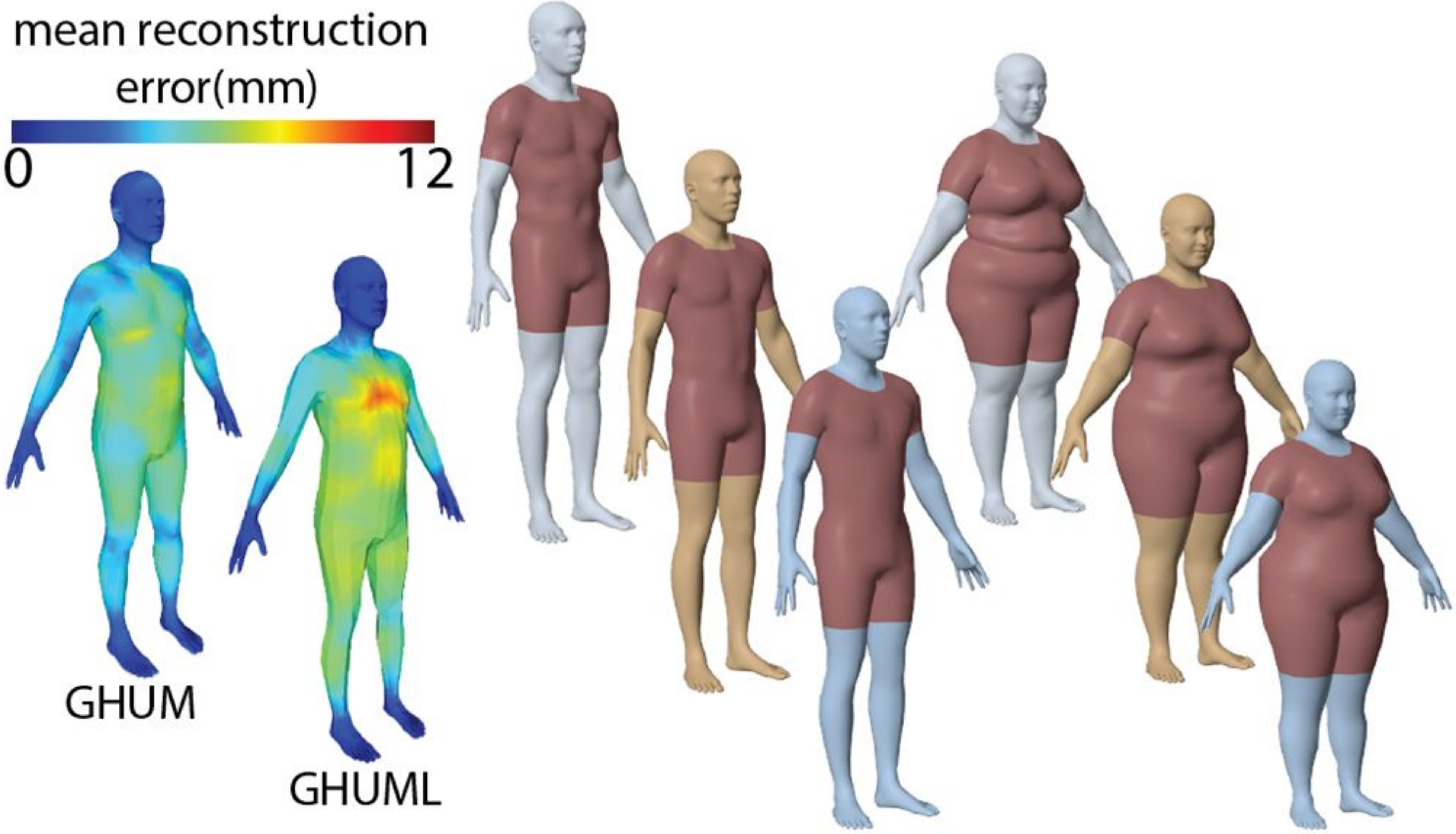


EVALUATION: CAESAR RECONSTRUCTION

mean vertex reconstruction error to registrations:

GHUM 2.81 mm

GHUML 3.27 mm



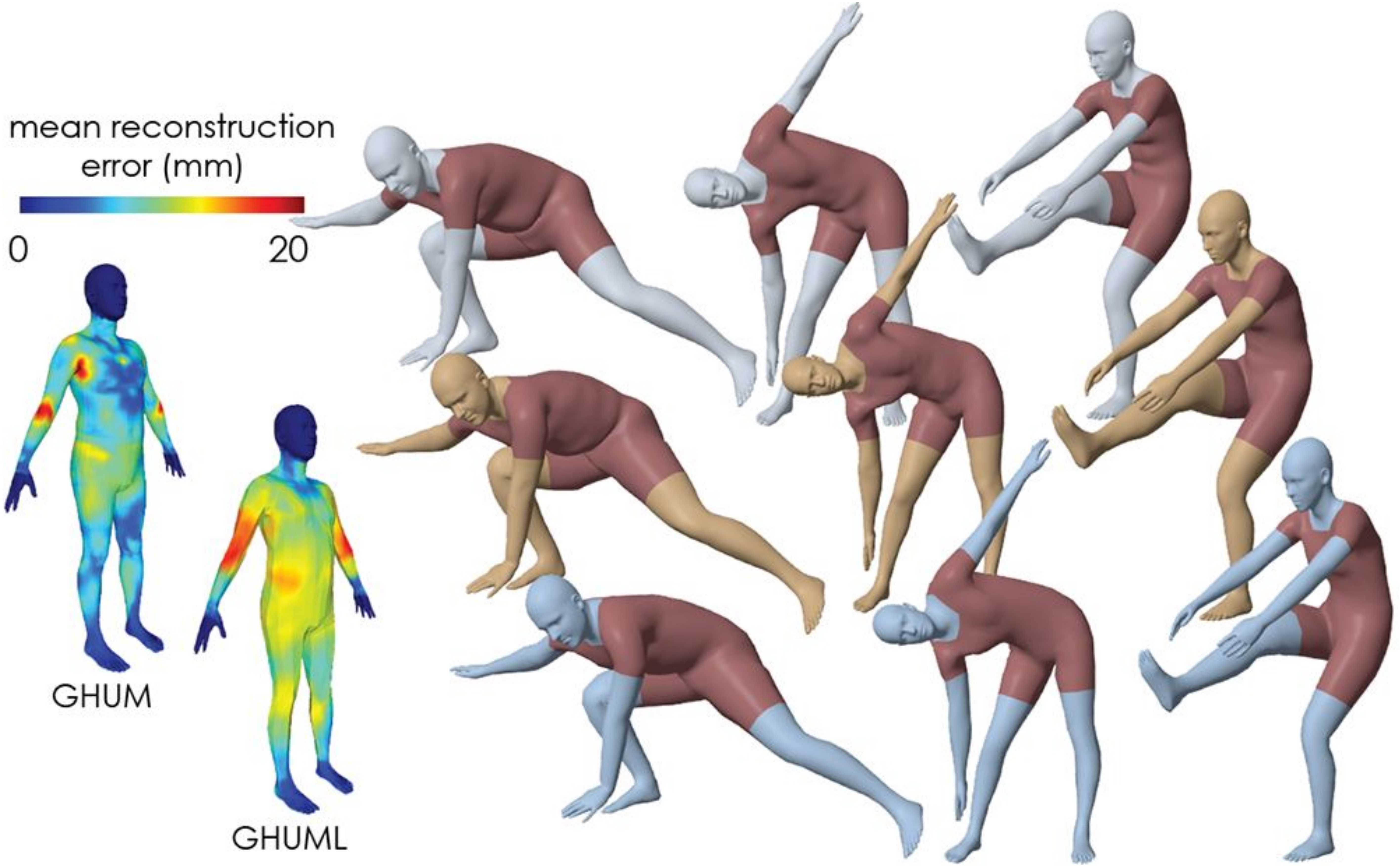
Google Research

EVALUATION: GHS3D BODY RECONSTRUCTION

mean vertex reconstruction error to registrations:

GHUM 5.21 mm

GHUML 6.32 mm

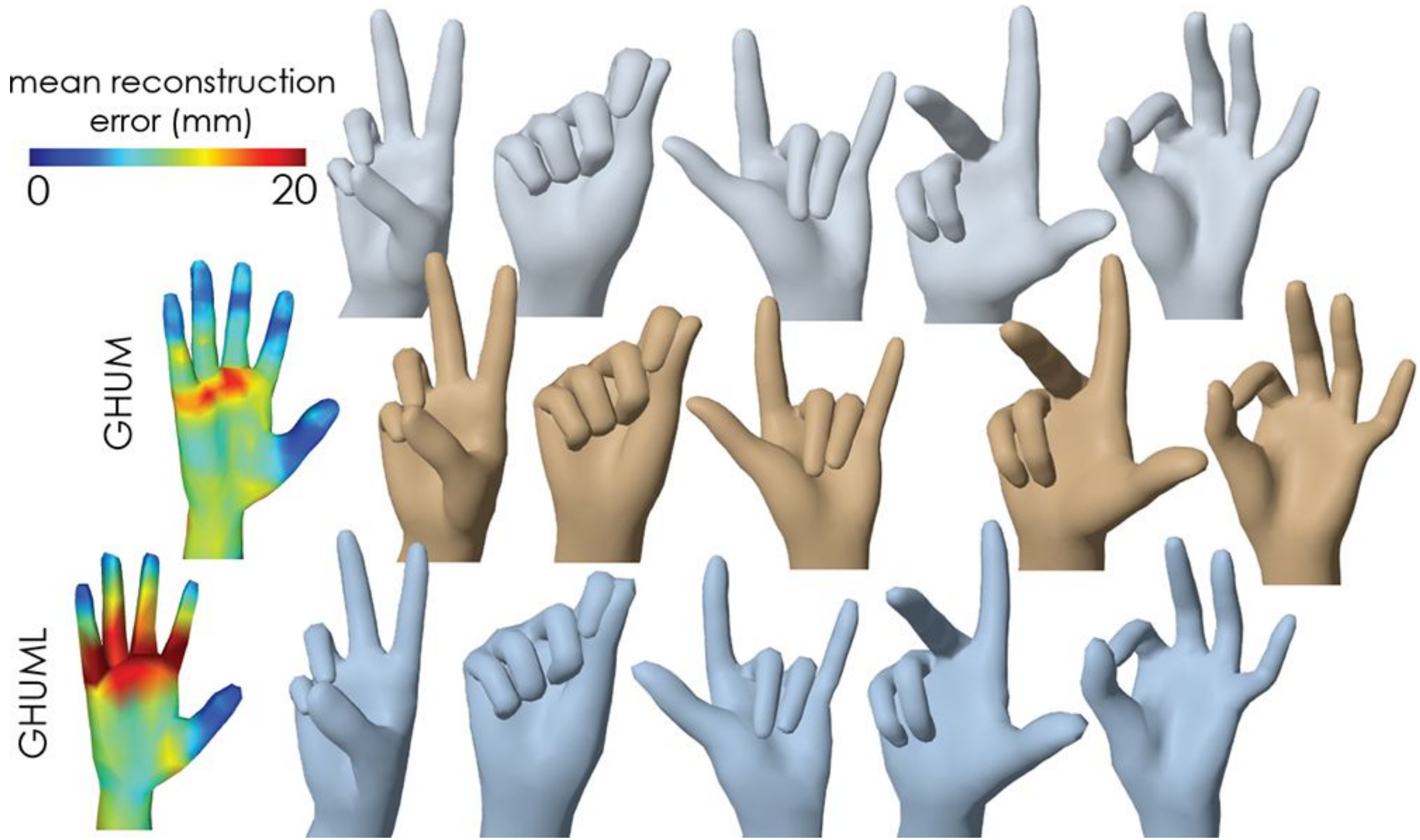


EVALUATION: GHS3D HAND RECONSTRUCTION

mean vertex reconstruction error to registrations:

GHUM 2.22 mm

GHUML 2.81 mm



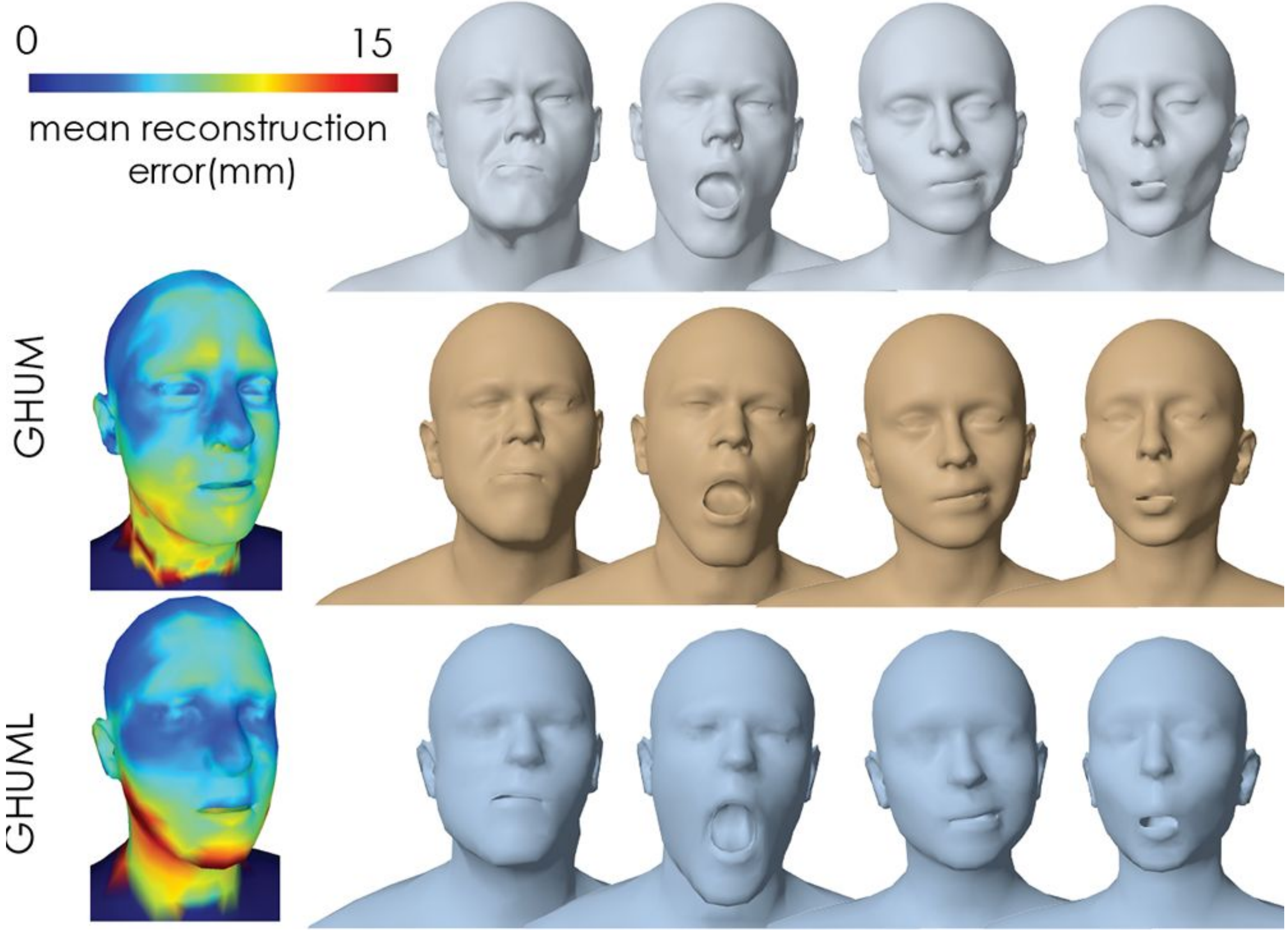
EVALUATION: GHS3D HEAD RECONSTRUCTION

Head motion =
joints (jaw, eyelids, neck) articulation +
facial expressions

mean vertex reconstruction error to registrations:

GHUM 2.96 mm

GHUML 3.82 mm



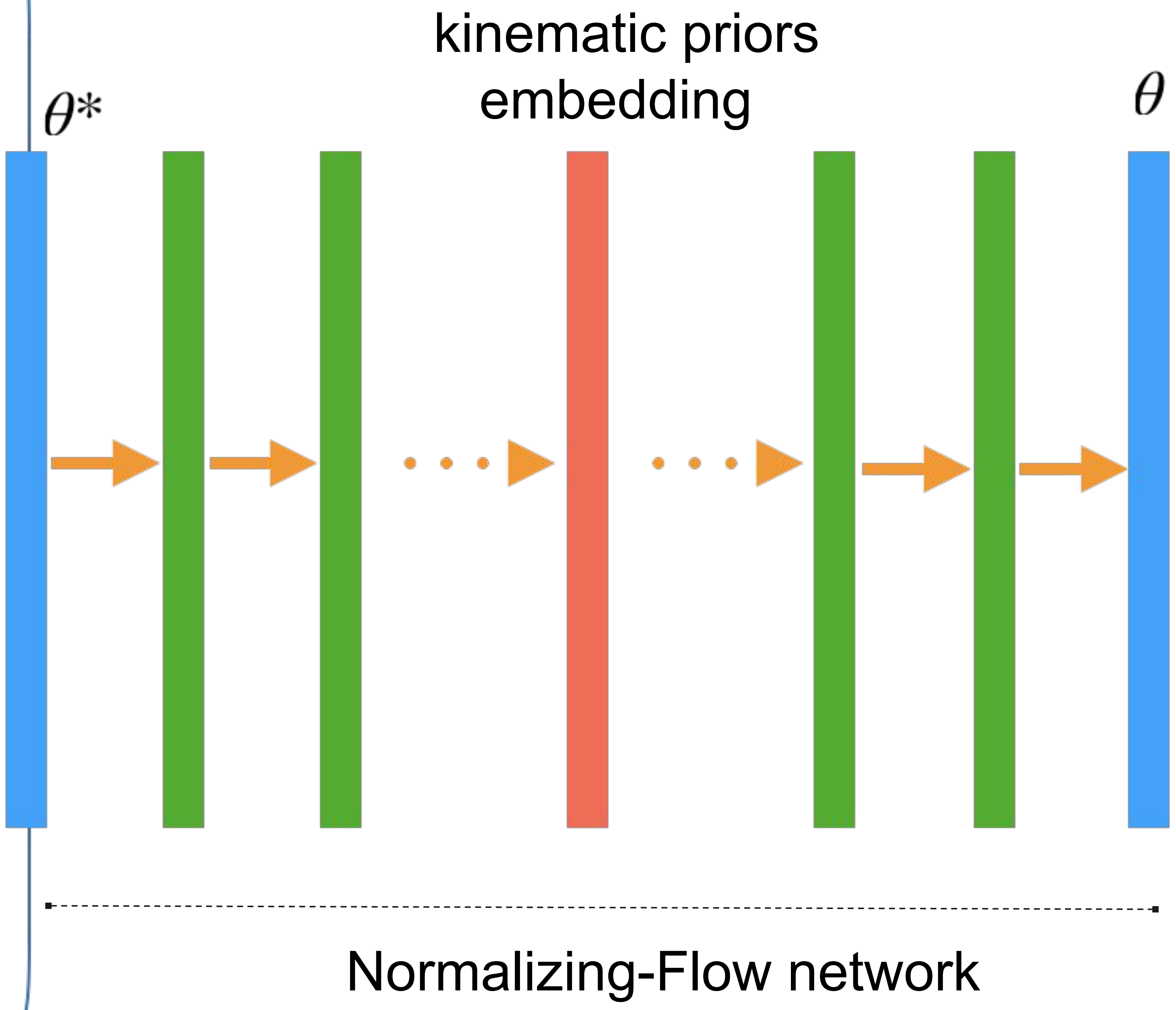
GHUM VS SMPL

- GHUM is close (or slightly better) to SMPL in skinning visual quality
- Vertex point-to-plane error (body-only)

  GHUM: 4.23 mm
  SMPL: 4.96 mm

MOTION RETARGETING AND KINEMATIC PRIORS

We retarget our models to ***2.8M*** CMU and ***2.2M*** Human36M motion capture frames.



FULL-BODY RECONSTRUCTION IN MONOCULAR IMAGES

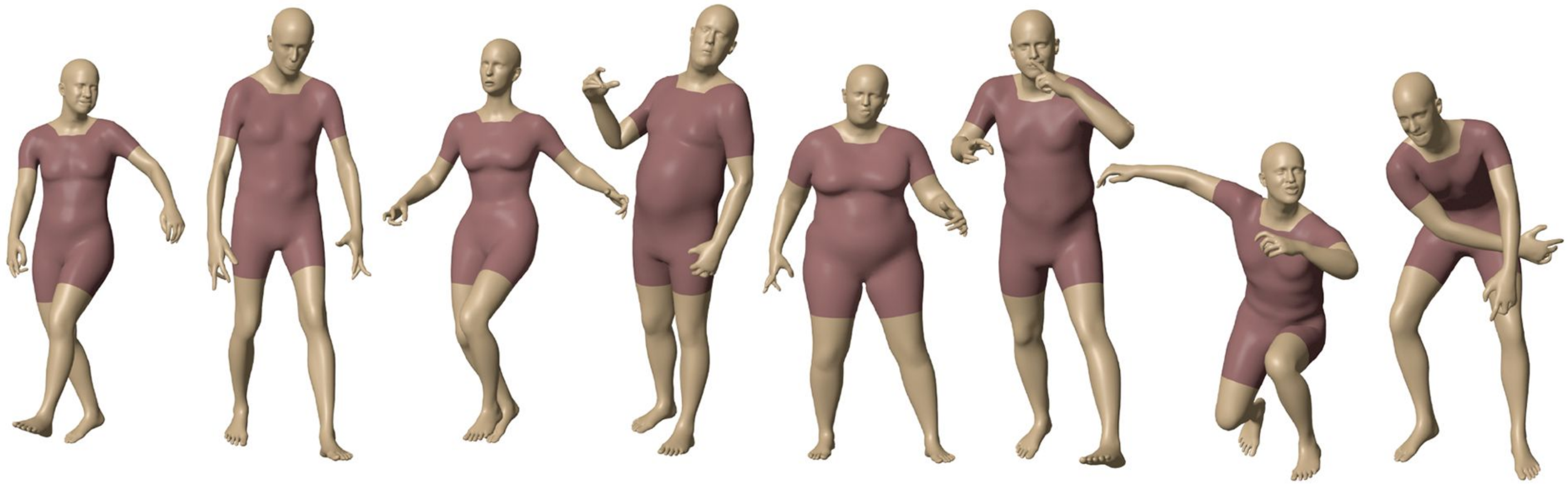
Google Research

[A.Zanfir et al. 2020] Weakly Supervised 3D Human Pose and Shape Reconstruction with Normalizing Flows

CONCLUSION

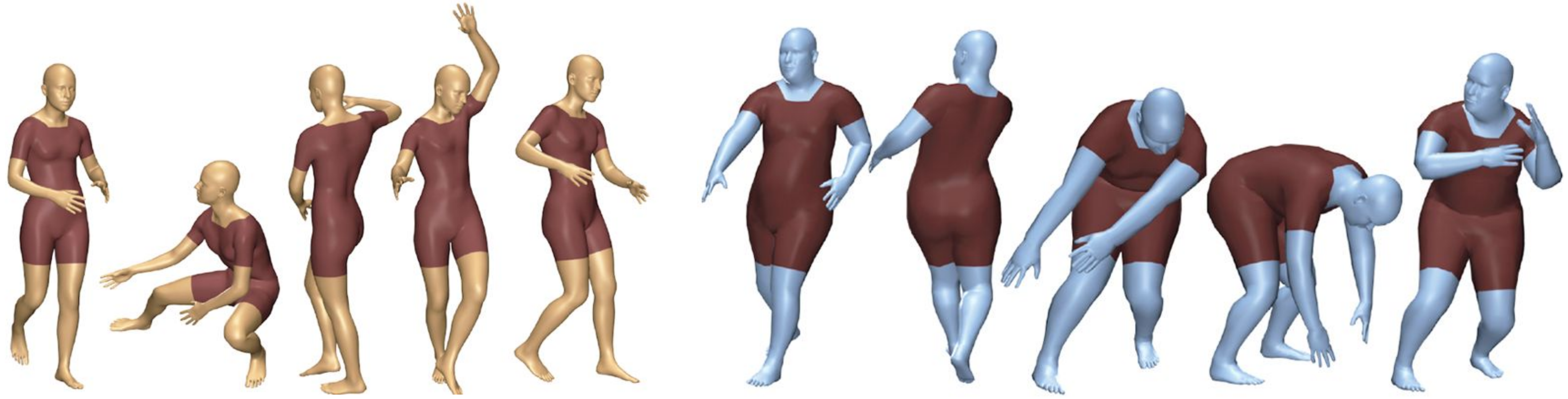
- An end-to-end generative, articulated 3D human shape learning pipeline.

- GHUM & GHUML: two high-quality full-body generative human models, available for research (https://github.com/google-research/google-research/ghum).



Thank You

Website: https://github.com/google-research/google-research/ghum



Google Research