

3D Scanning & Motion Capture

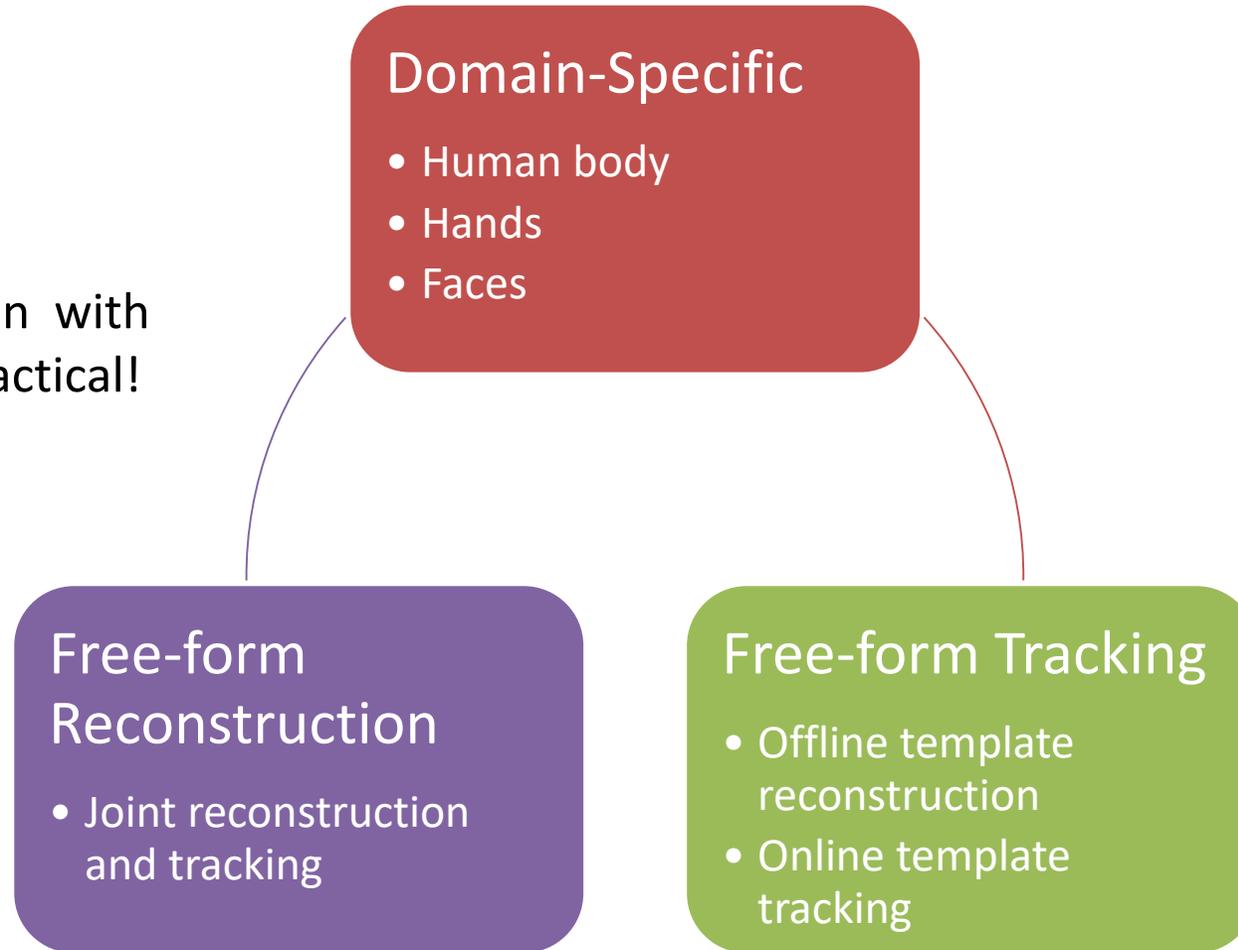
Body and Hand Tracking

Prof. Niessner

tum:3D

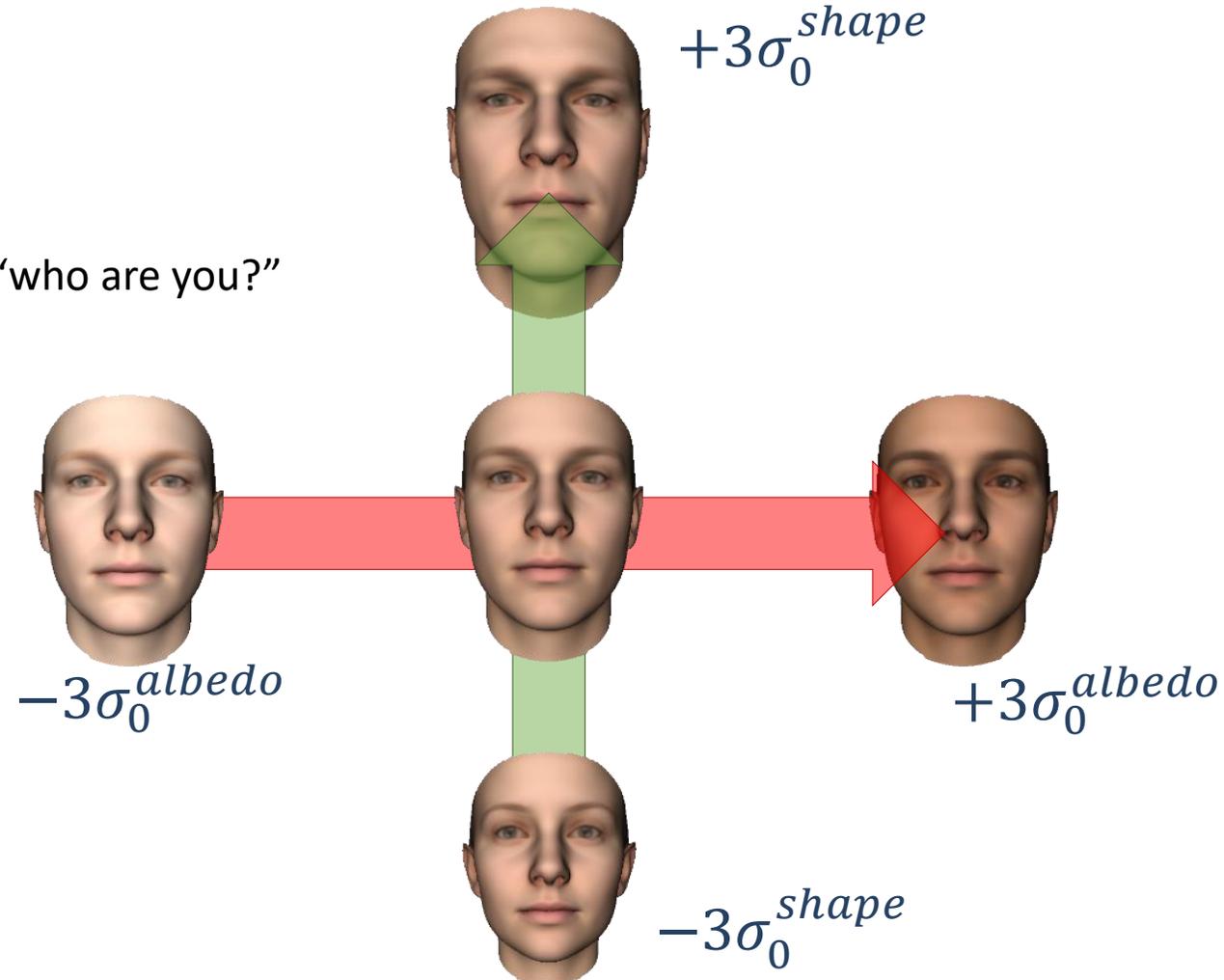
Dynamic 3D Capture

Need more regularization with fewer DoF to make it practical!



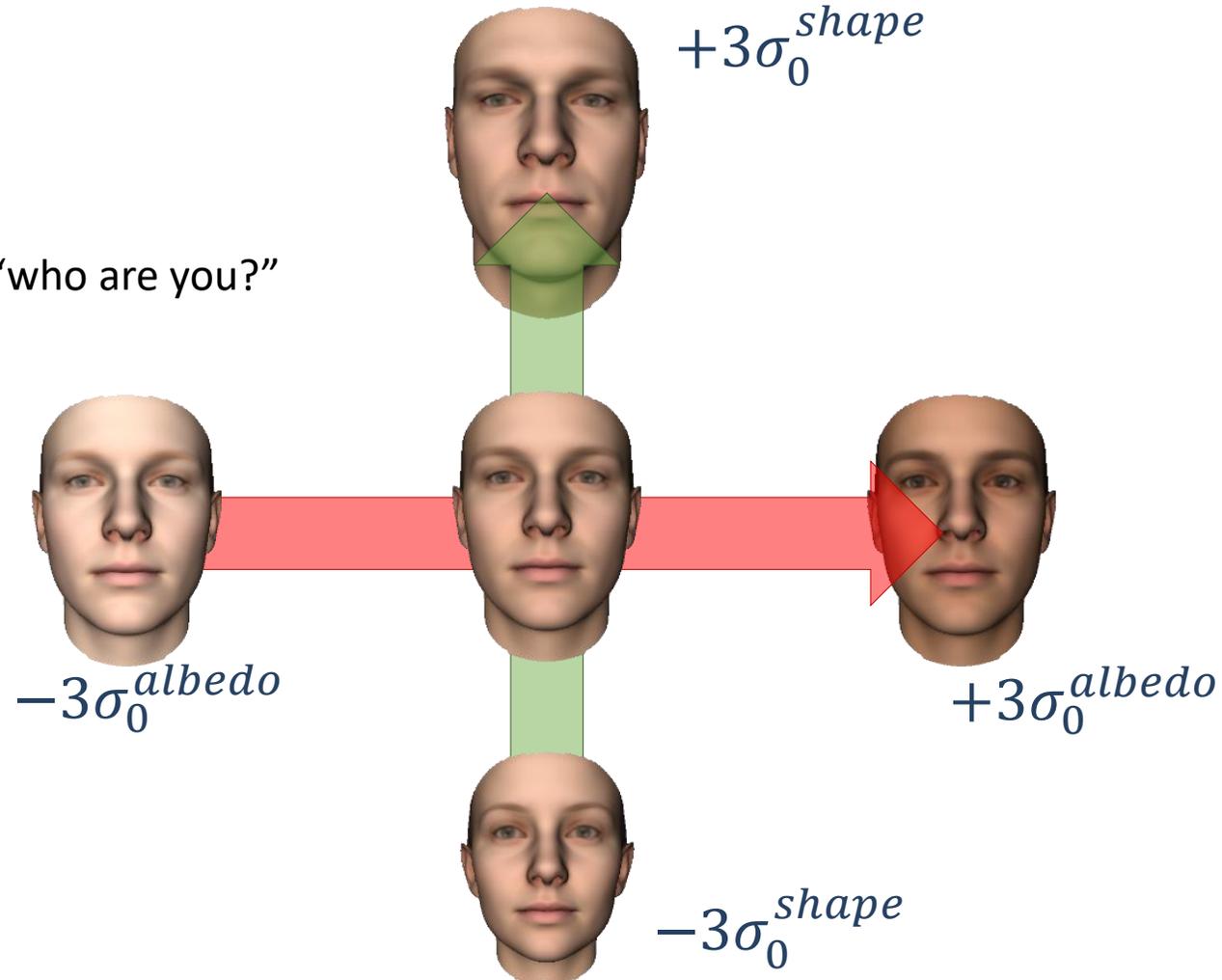
Parametric Face Model: Shape Identity

Shape Identity defines “who are you?”



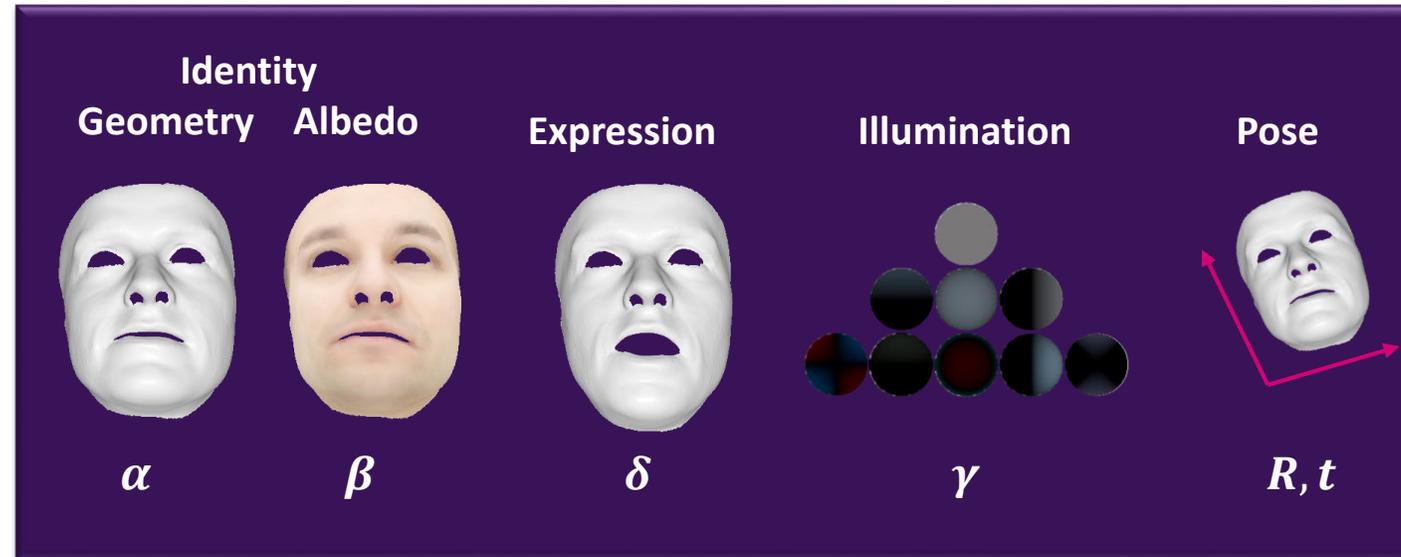
Parametric Face Model: Shape Identity

Shape Identity defines “who are you?”



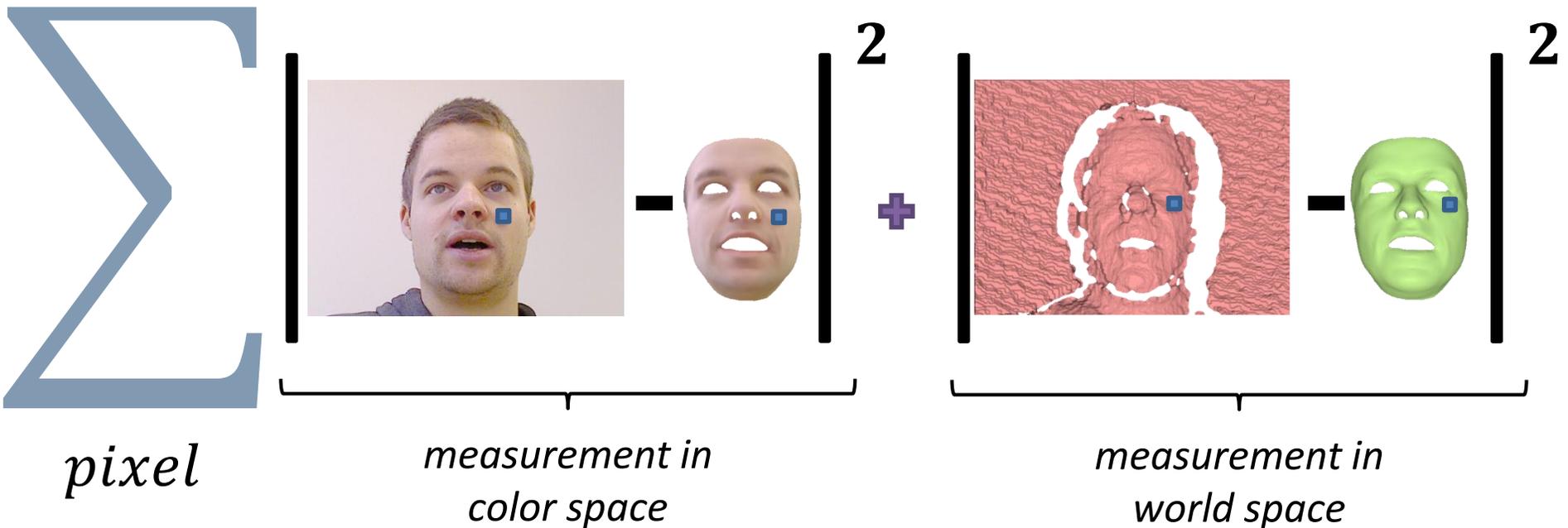
Face Model Fitting

$$P =$$



Parameter Estimation as Energy Minimization

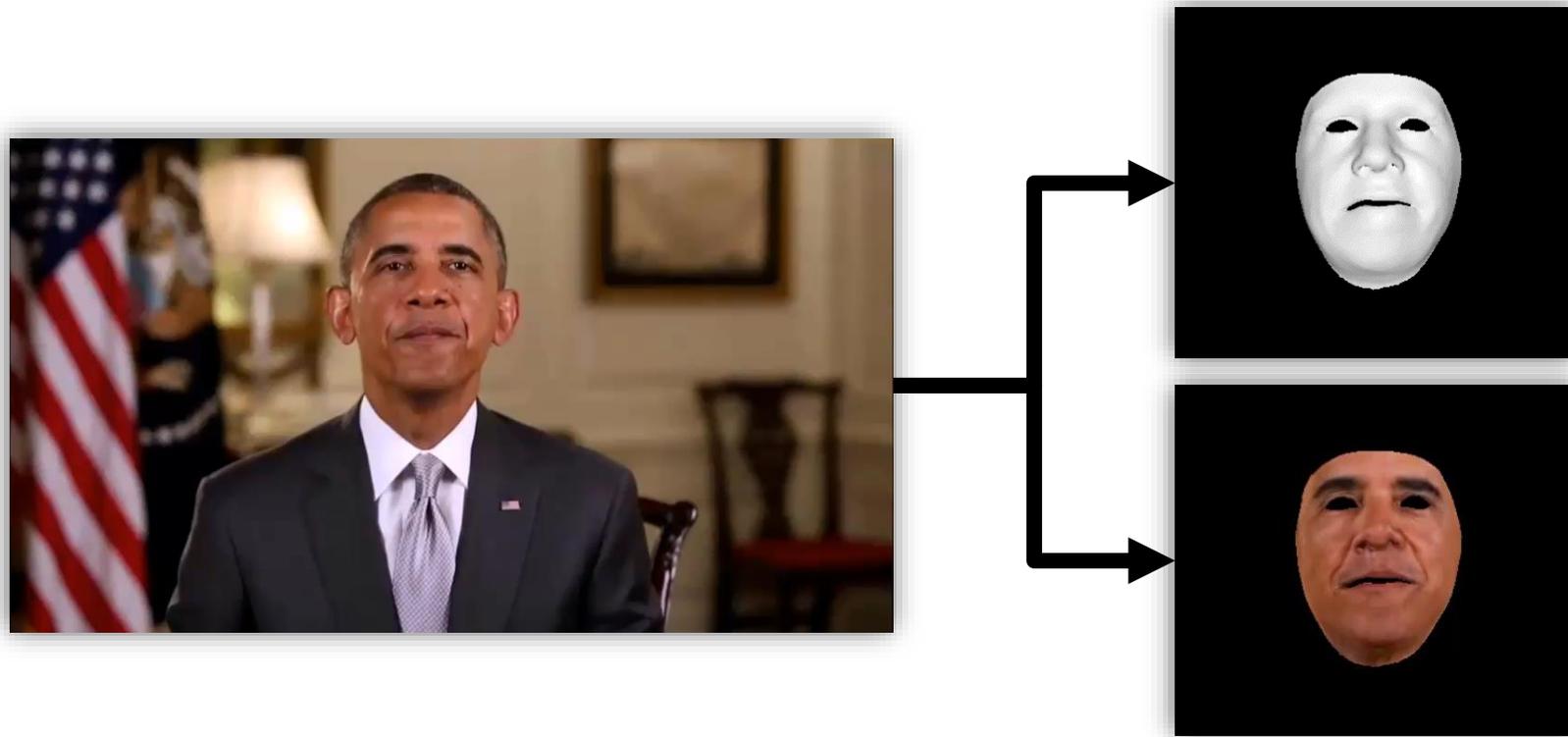
$$E(P) = \underbrace{E_{dense}(P)}_{data} + E_{sparse}(P) + \underbrace{E_{reg}(P)}_{prior}$$

$$E_{dense}(P) = \sum_{pixel} \left[\left| \text{measurement in color space} \right|^2 + \left| \text{measurement in world space} \right|^2 \right]$$


The diagram illustrates the dense energy term $E_{dense}(P)$ as a summation over pixels. It shows two squared differences being summed:

- measurement in color space:** A color image of a man's face minus a white 3D mesh of the same face.
- measurement in world space:** A red depth map of the same face minus a green 3D mesh of the same face.

RGB-only Face Tracking

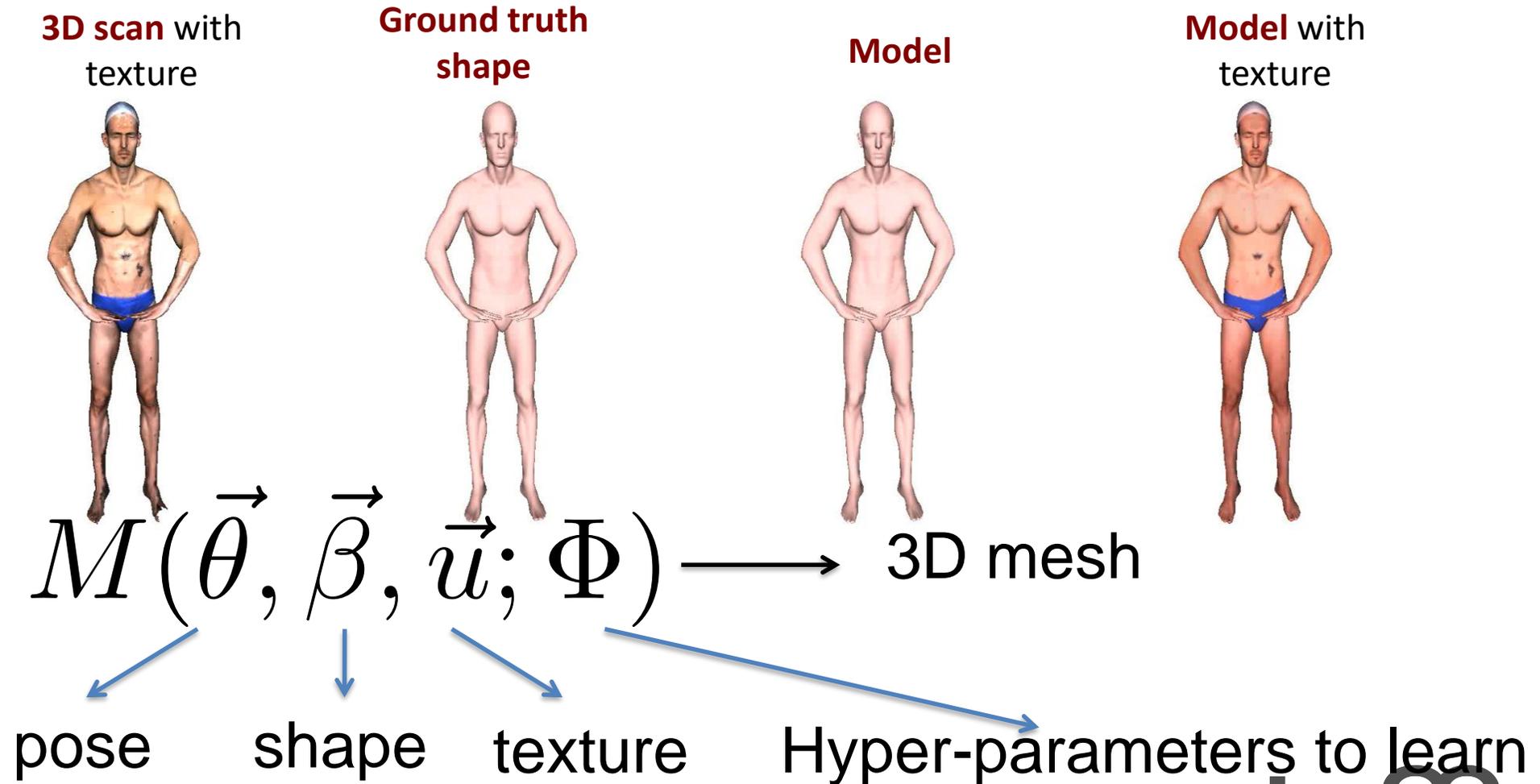


Today: Body and Hand Tracking

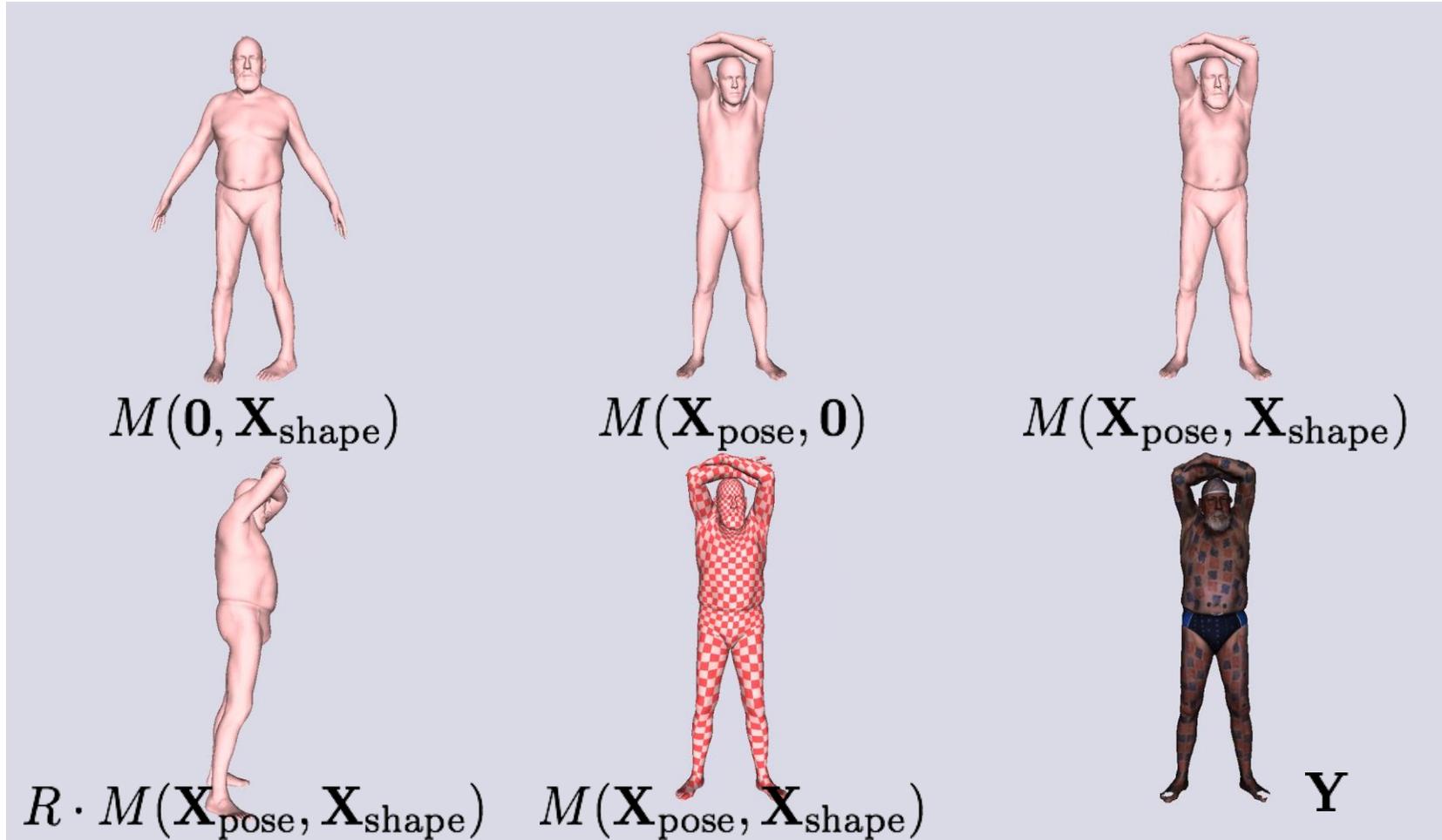
Parametric Models

- Idea of parametric generalizes to other parts: hands and bodies
- Simple roadmap
 - Take set of scans from variety of subjects / poses
 - Non-rigidly align them with template to obtain consistent topology
- Often parameterization with respect to an underlying skeleton

Parametric Models: Virtual Humans

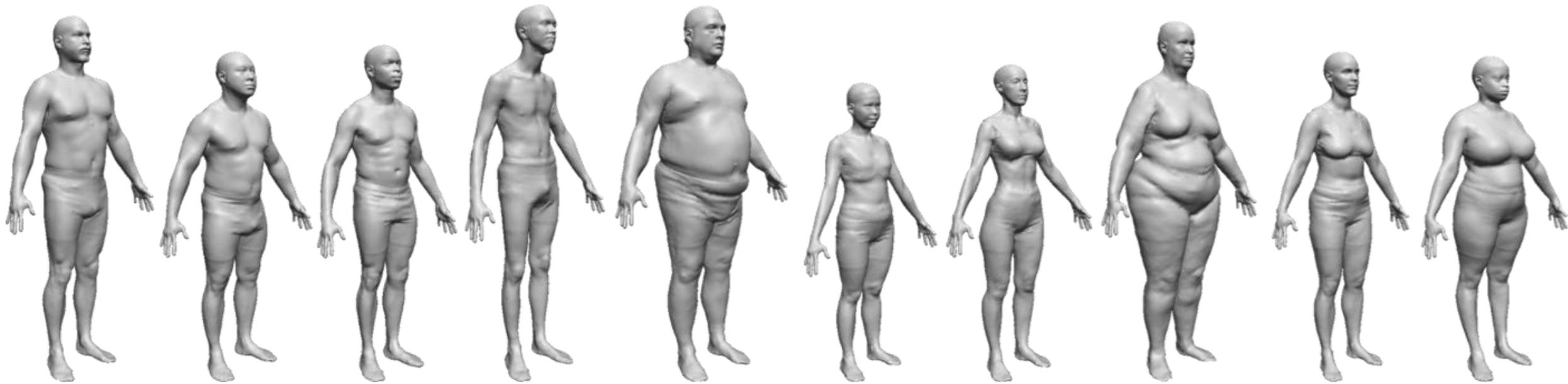


A Body Model is a Function



Modeling the Space of Human Bodies

- CAESAR dataset (Civilian American and European Surface Anthropometry Resource Project)

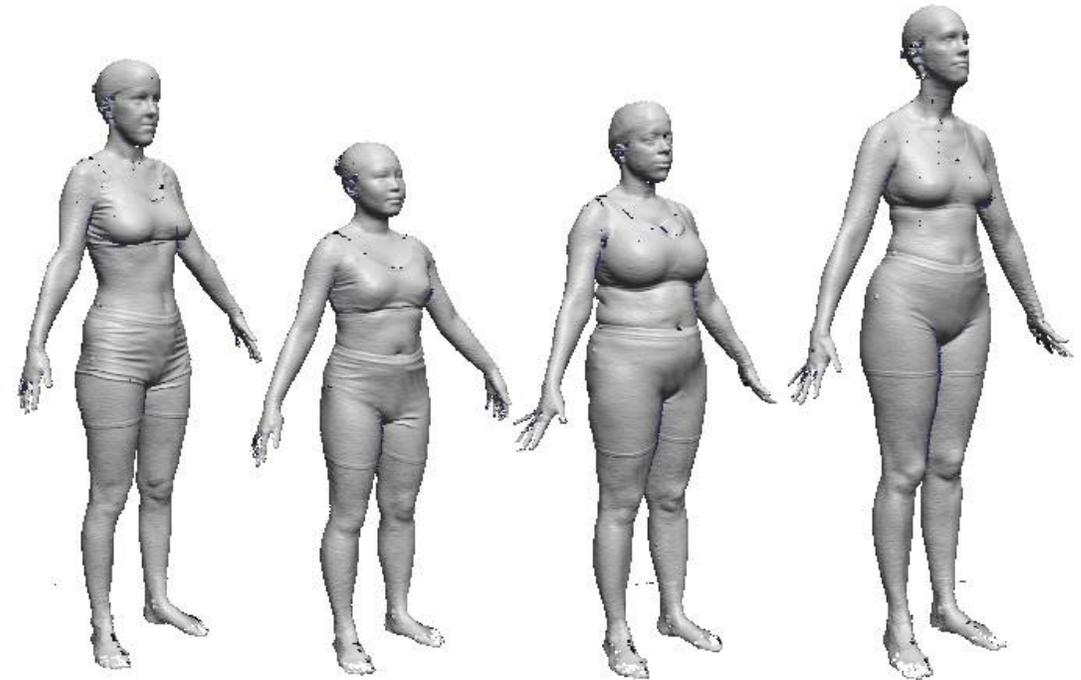


Modeling the Space of Human Bodies

- CAESAR dataset (Civilian American and European Surface Anthropometry Resource Project)

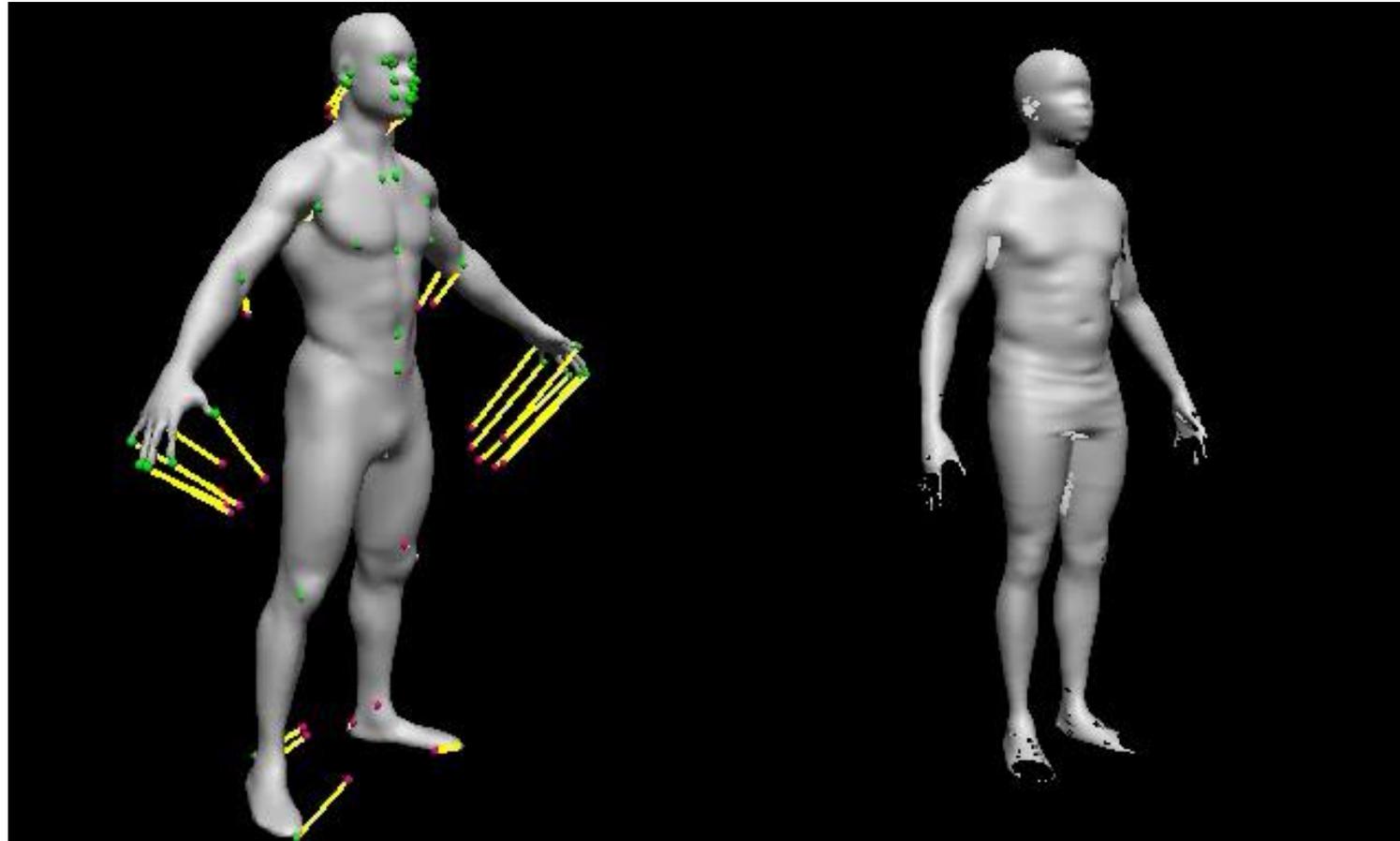


Cyberware 3D whole body scanner



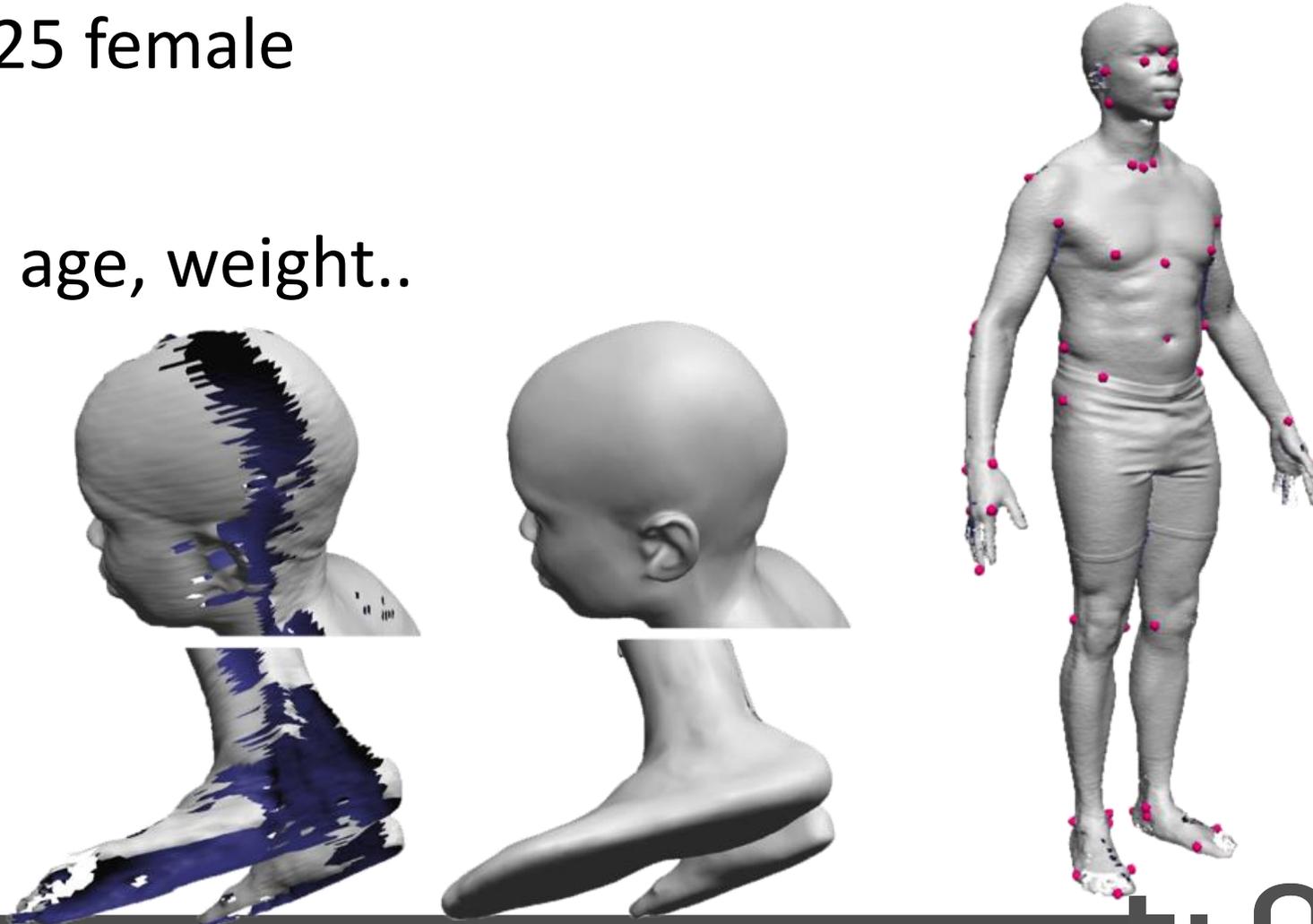
Modeling the Space of Human Bodies

- Non-rigid registration to align scans



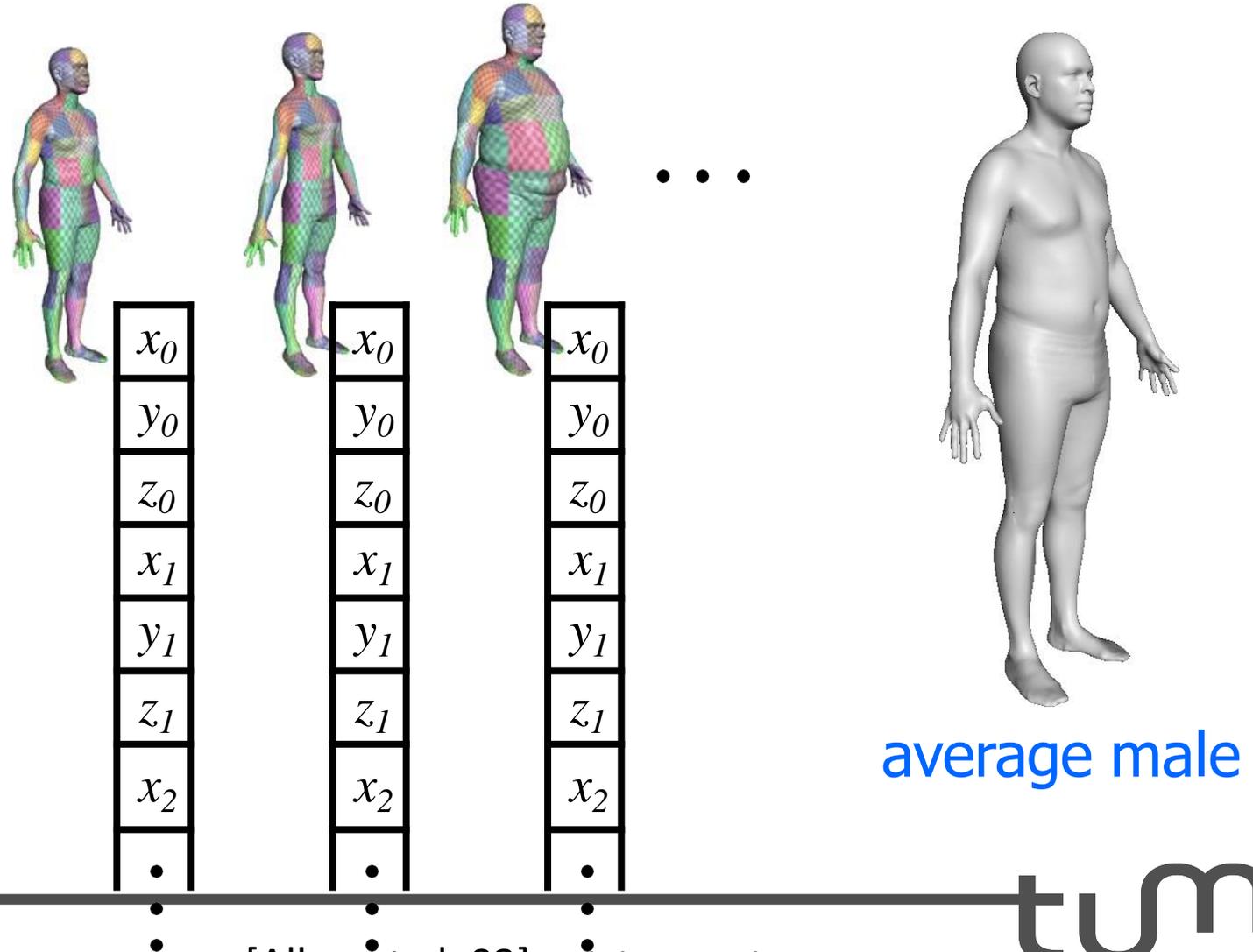
Modeling the Space of Human Bodies

- 125 male and 125 female
- 74 markers
- Various ethnics, age, weight..

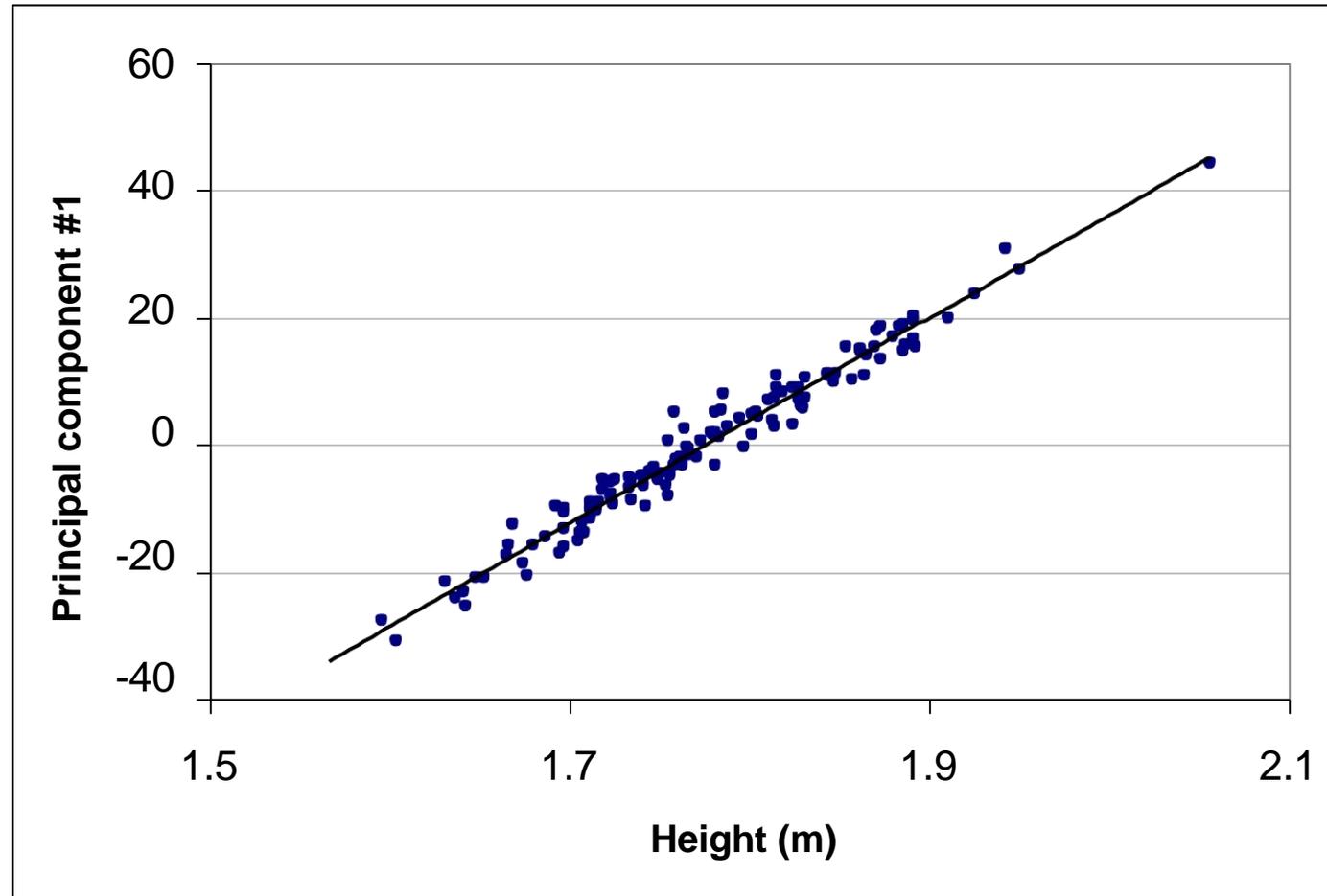


Variance in Human Bodies

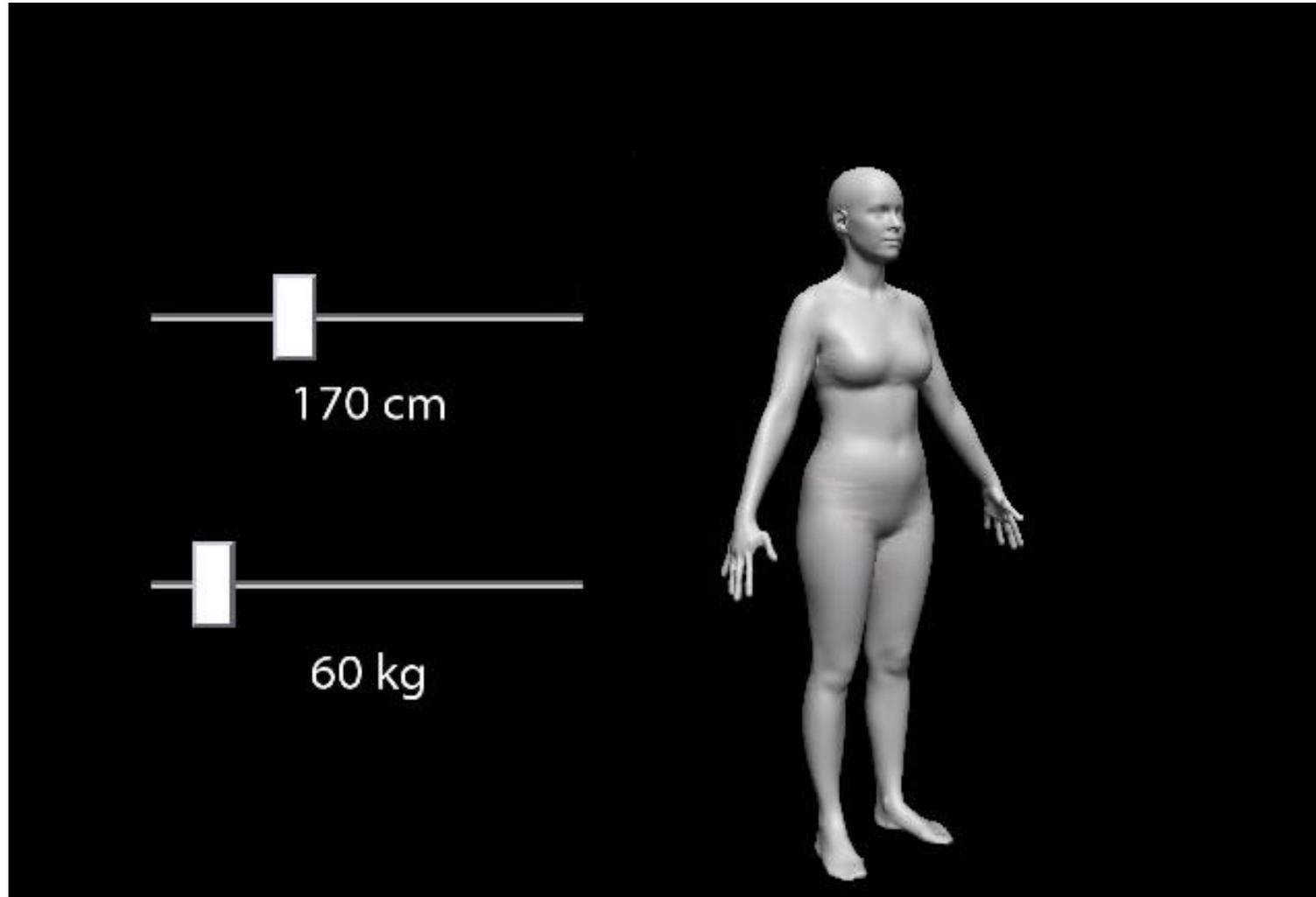
Compute statistical model using PCA



Variance in Human Bodies

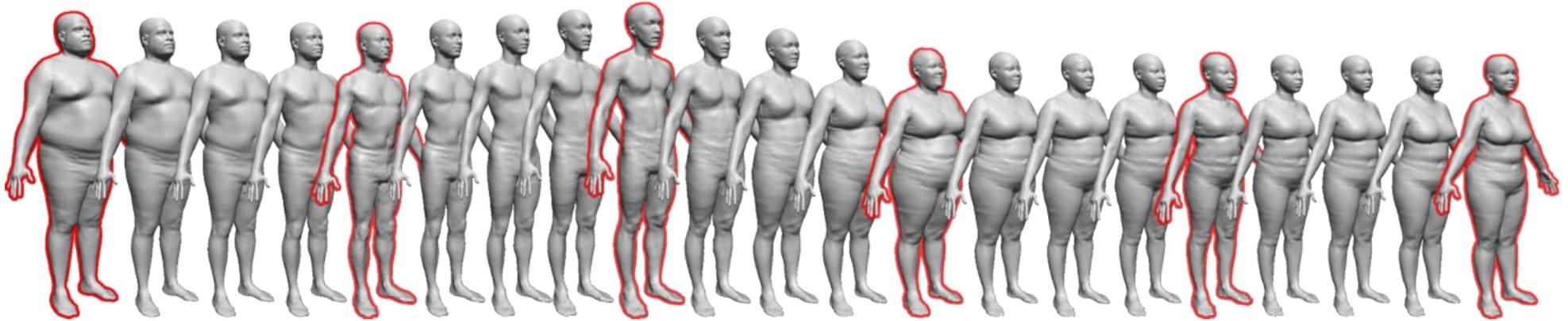


Variance in Human Bodies

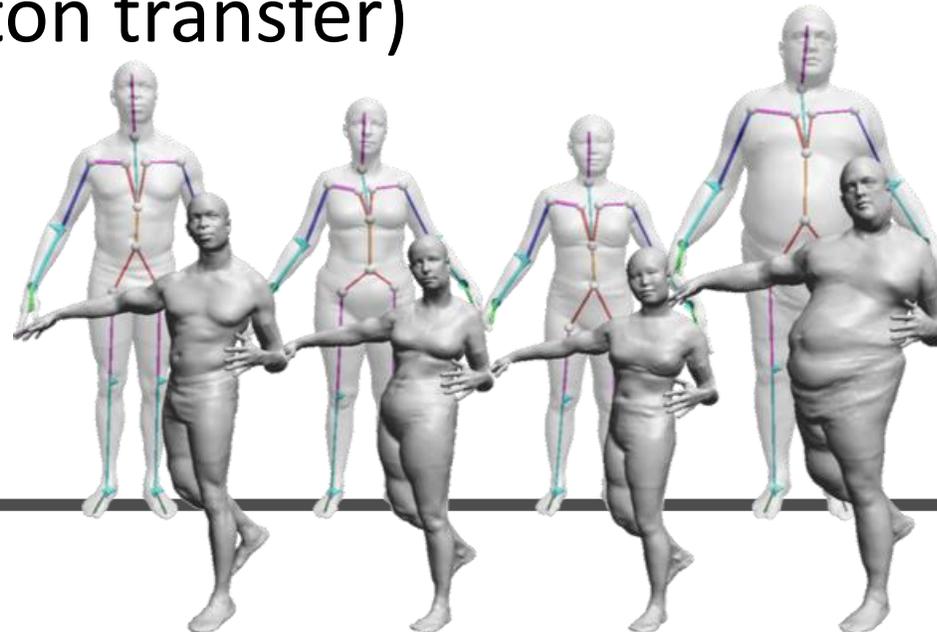


Variance in Human Bodies

- Morphing



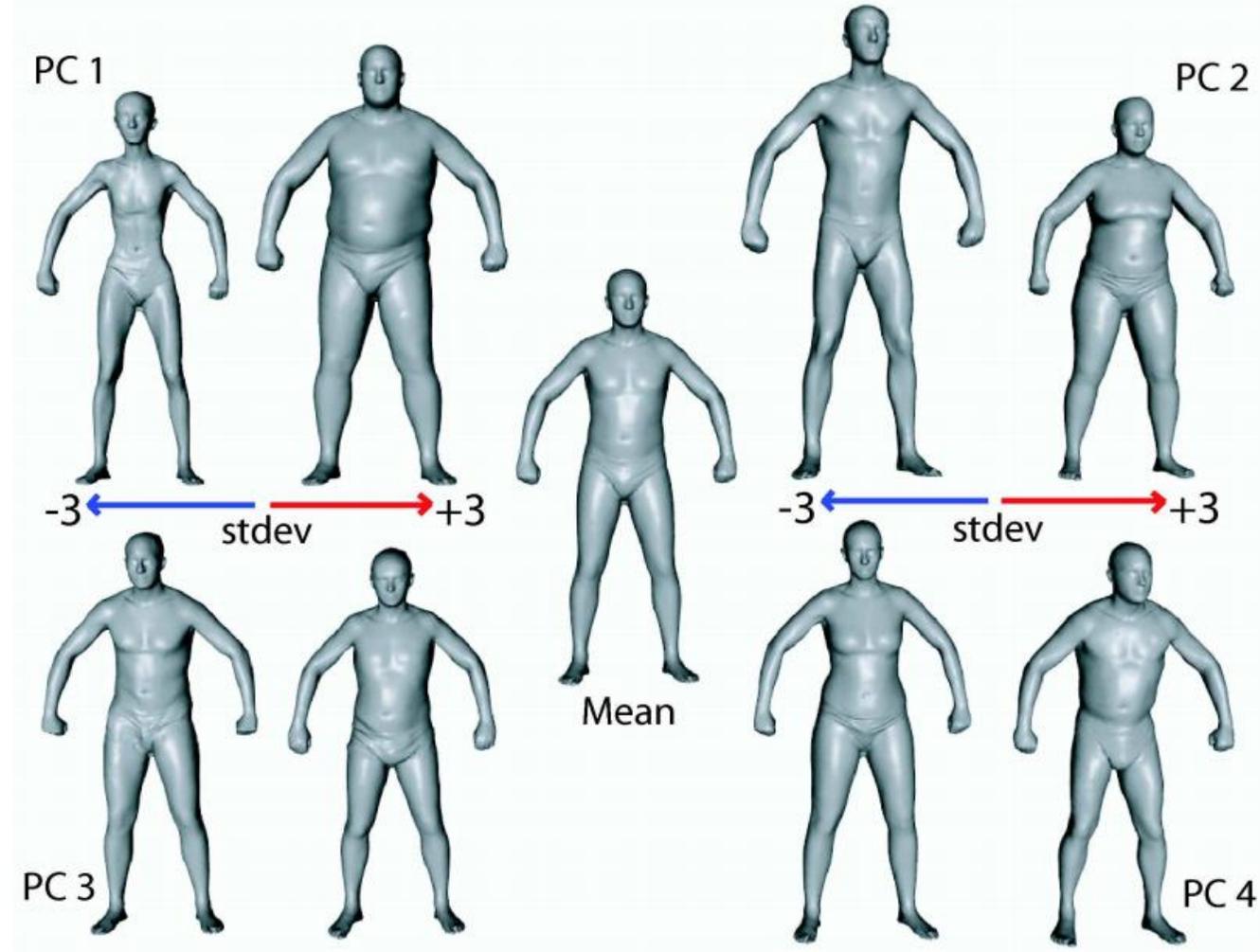
- Retargeting (skeleton transfer)



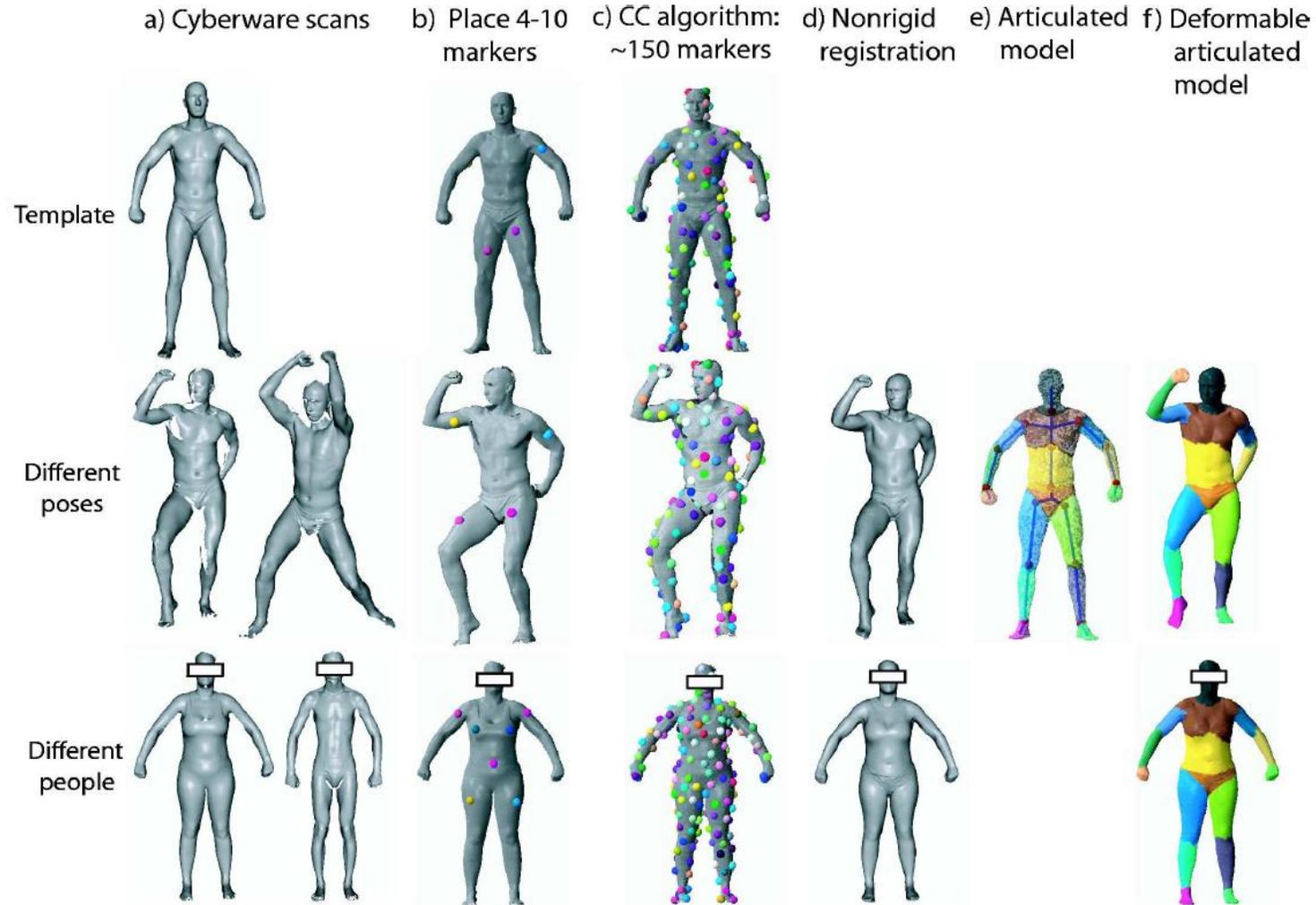
SCAPE: Shape Completion and Animation of People

- Cyberware scanner: captures four directions simultaneously; approx. 200k points / model; merging four scan views
- Poses: 70 poses of a particular person in a wide variety of poses
- Identity: 37 different people in a similar (but not identical) pose
+ CEASAR
- Reconstruct skeleton (16 parts) for each pose
- Compute Delta-‘PCA’-like blendshapes over triangles wrt. skel.

Principle Components of SCAPE



SCAPE: Shape Completion and Animation of People



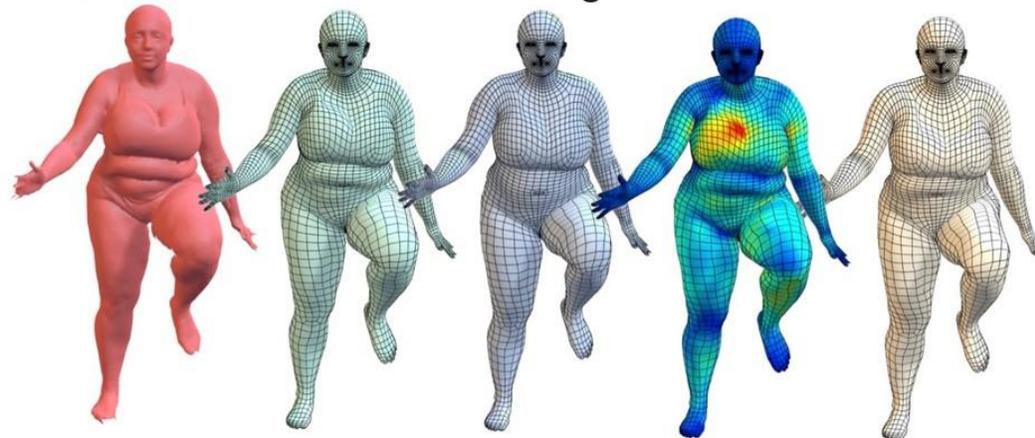
SMPL et al. (e.g., A Skinned Multi-Person Linear Model)



Breathing model



4D scanner (60 fps)



4D scan

4Cap

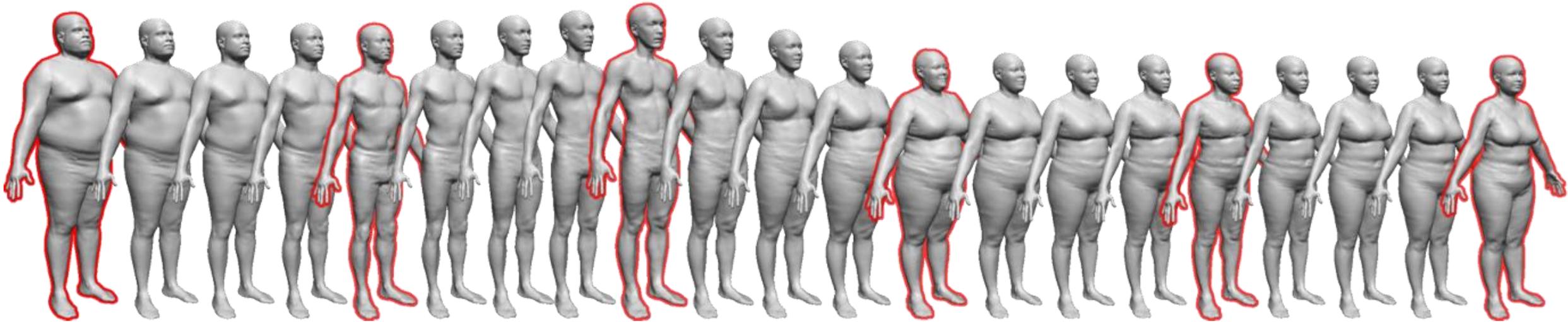
BlendSCAPE

Soft tissue

Dyna

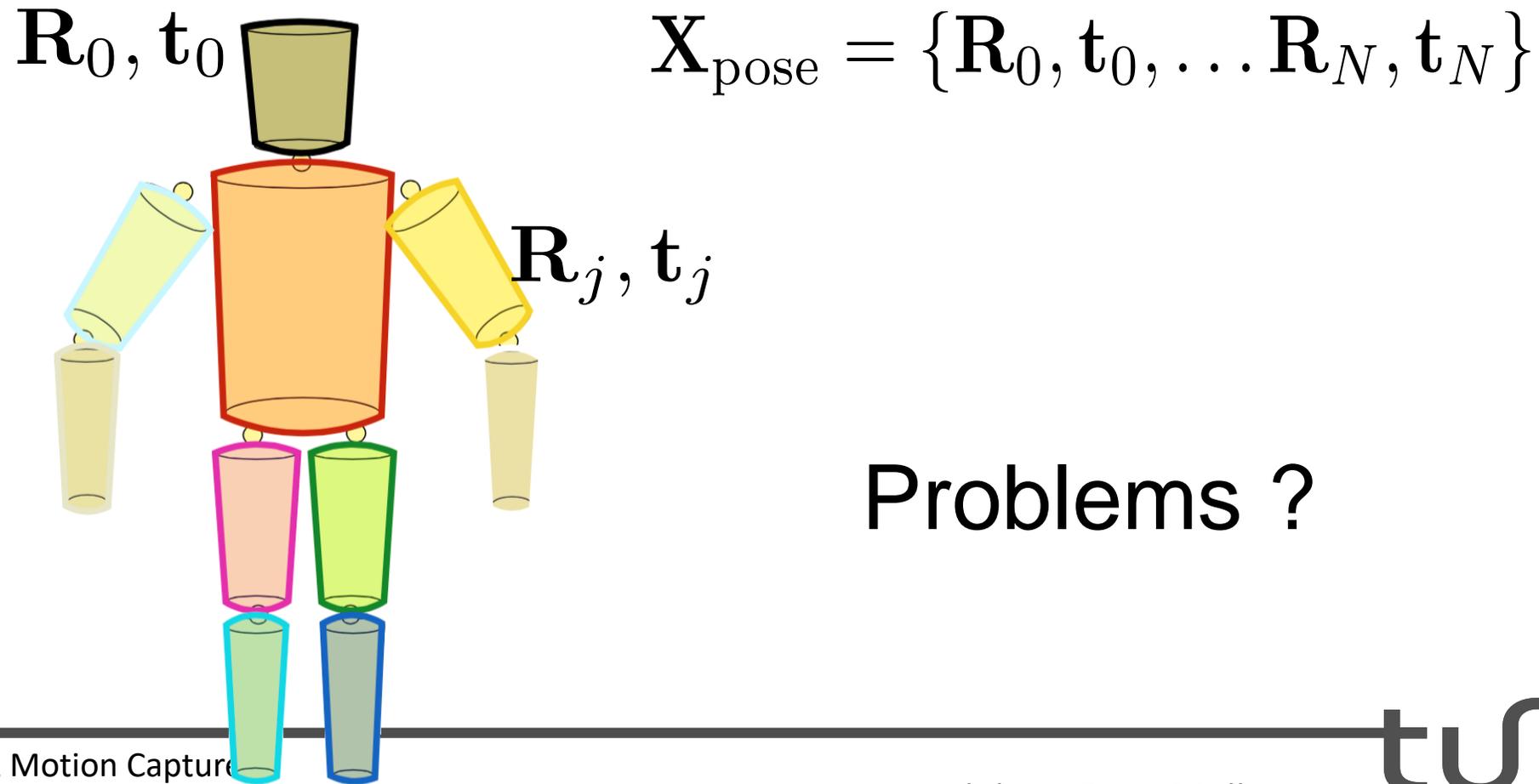
Parametrization of Shape

- PCA / Blendshapes



How do we parameterize pose ?

- Parameterize every body part separately ?



Problems ?

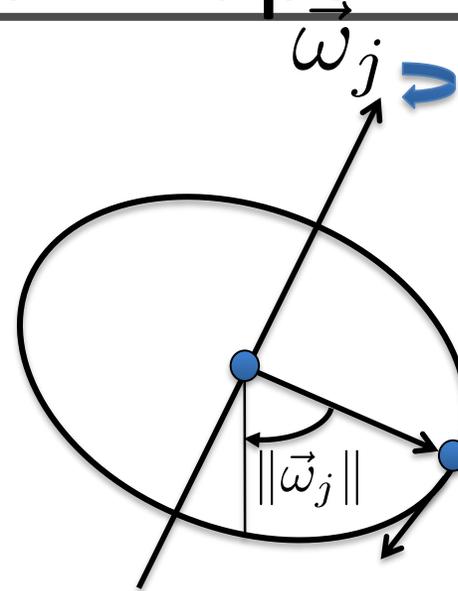
Rotation Parameterization

- Rotations are composed of 9 numbers
- **6 additional constraints** to ensure that the matrix is orthonormal
- **Suboptimal** for optimization

Rotation with Exponential Maps

$\|\vec{\omega}_j\|$: Angle of rotation

$\vec{\omega}_j$: scaled axis of rotation

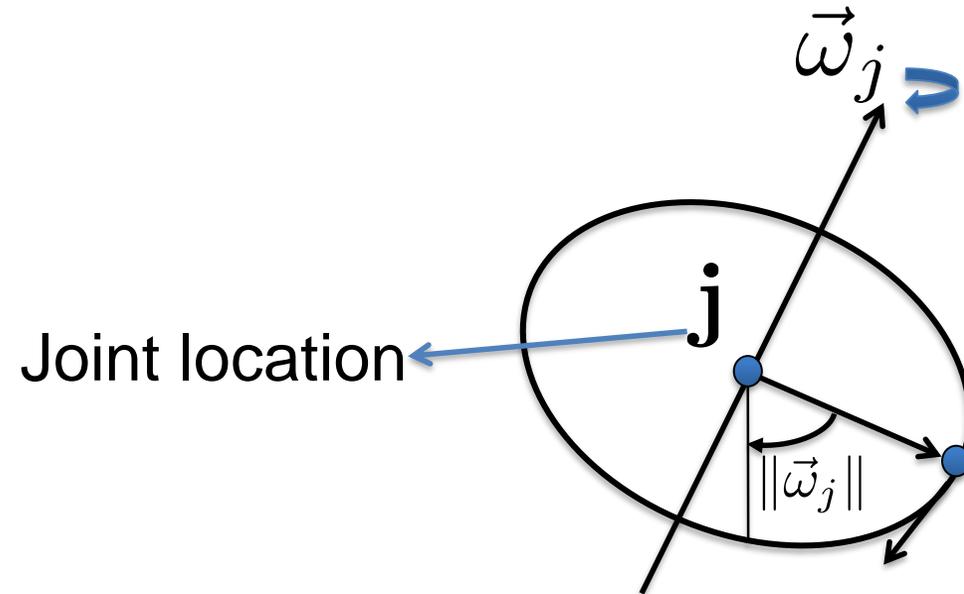


Rotation obtained with Rodrigues formula:

$$\mathbf{R} = e^{\hat{\omega}} = \mathcal{I} + \hat{\omega}_j \sin(\|\vec{\omega}_j\|) + \hat{\omega}_j^2 (1 - \cos(\|\vec{\omega}_j\|))$$

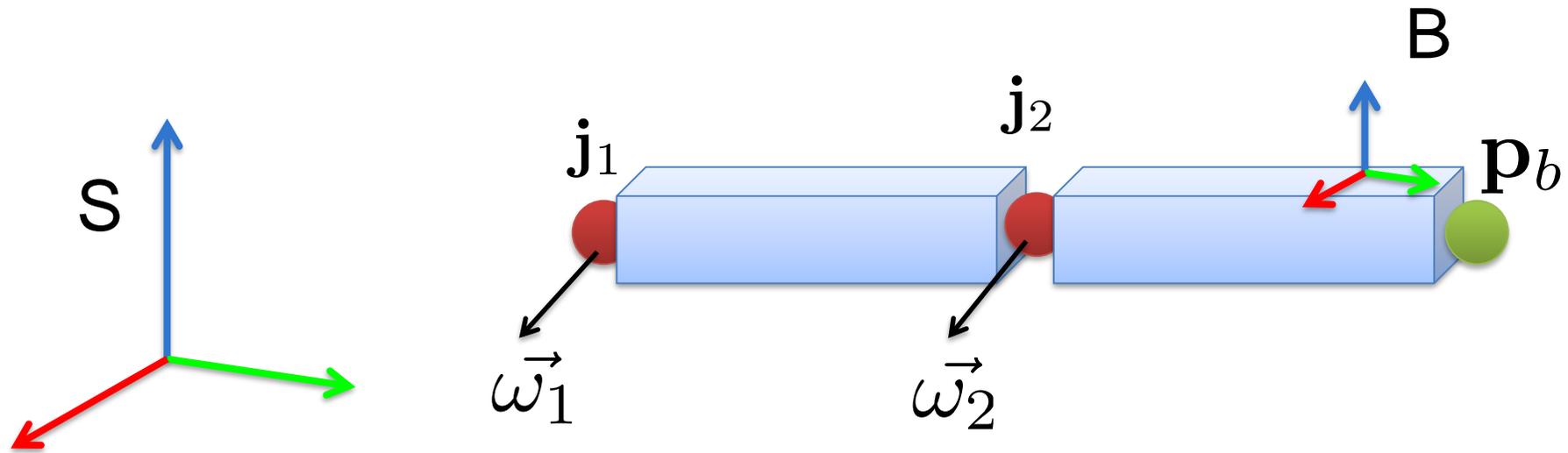
Joint Rigid Body Motion

The transformation associated with a rotational joint is

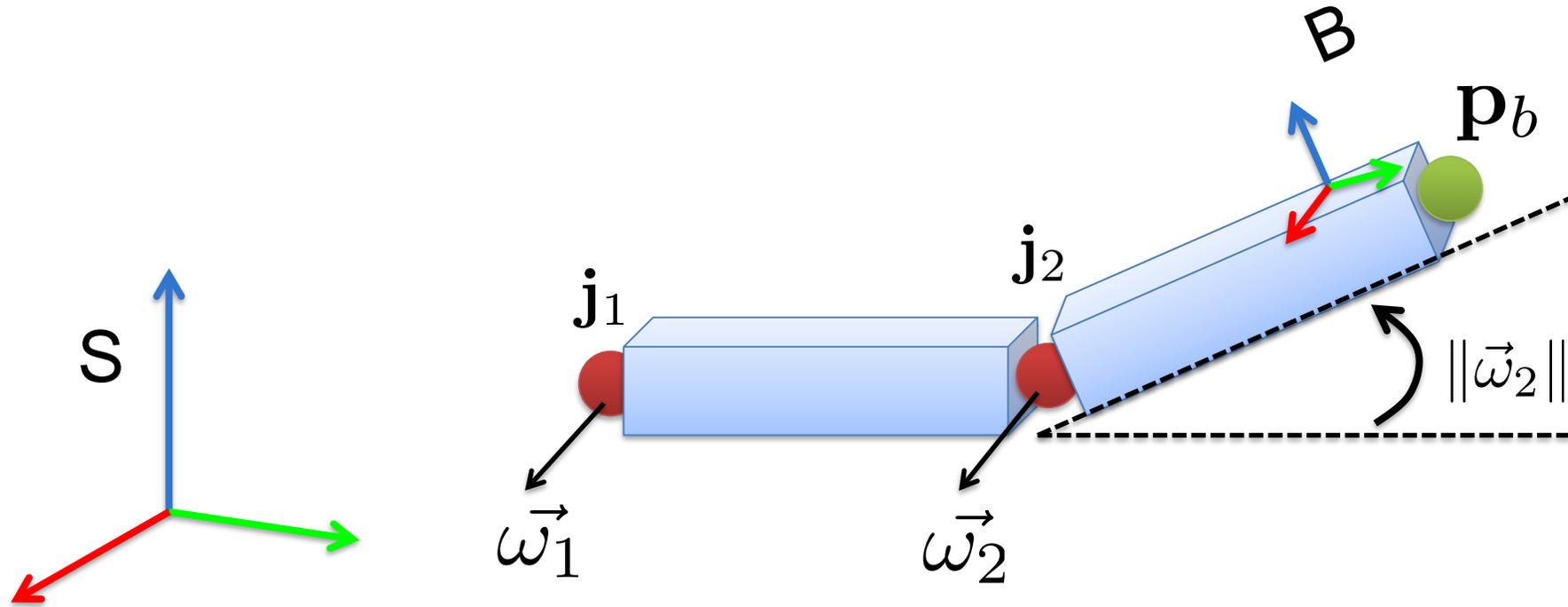


$$G(\vec{\omega}, \mathbf{j}) = \begin{bmatrix} [e^{\vec{\omega}}]_{3 \times 3} & \mathbf{j}_{3 \times 1} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix} \rightarrow \text{Rigid Body Motion}$$

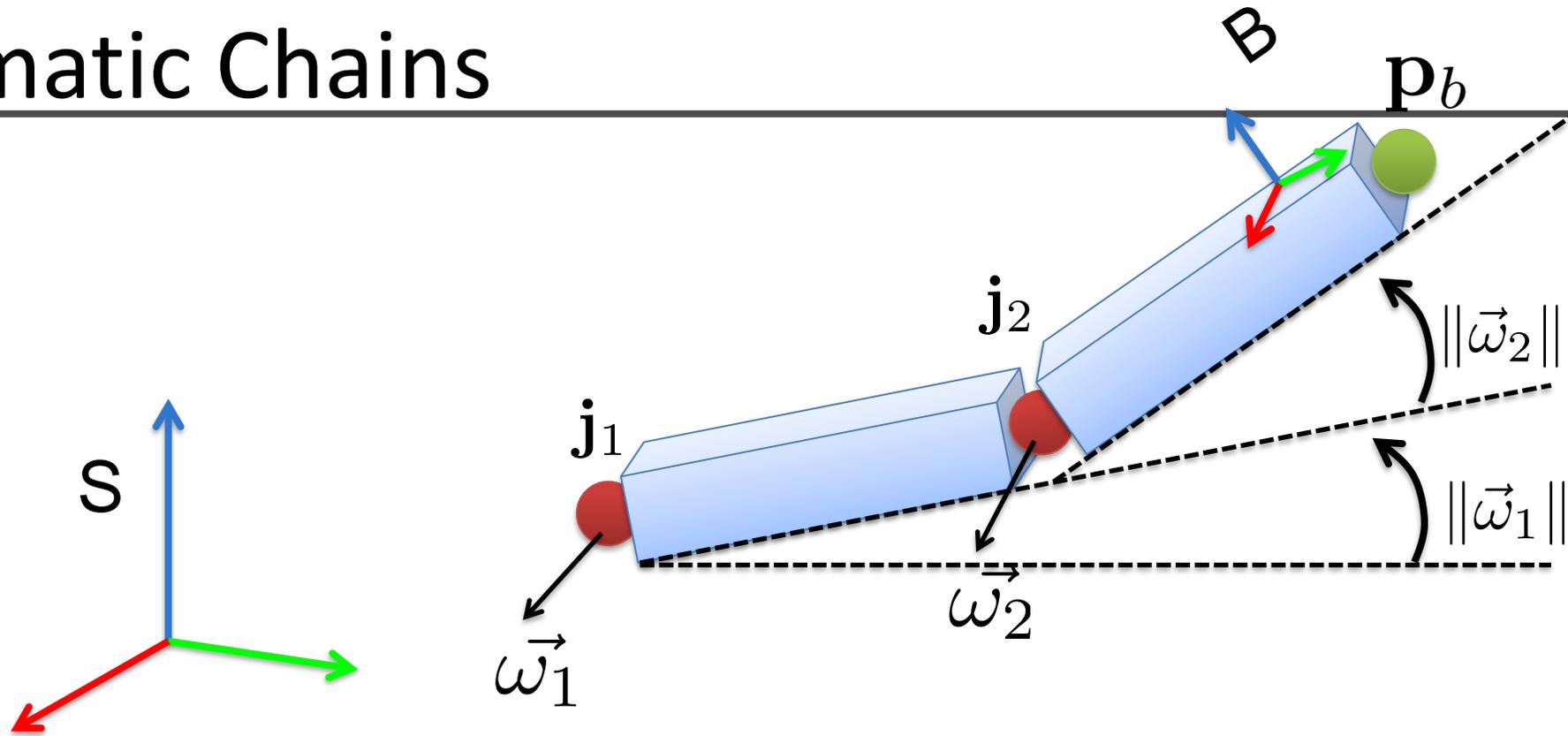
Kinematic Chains



Kinematic Chains



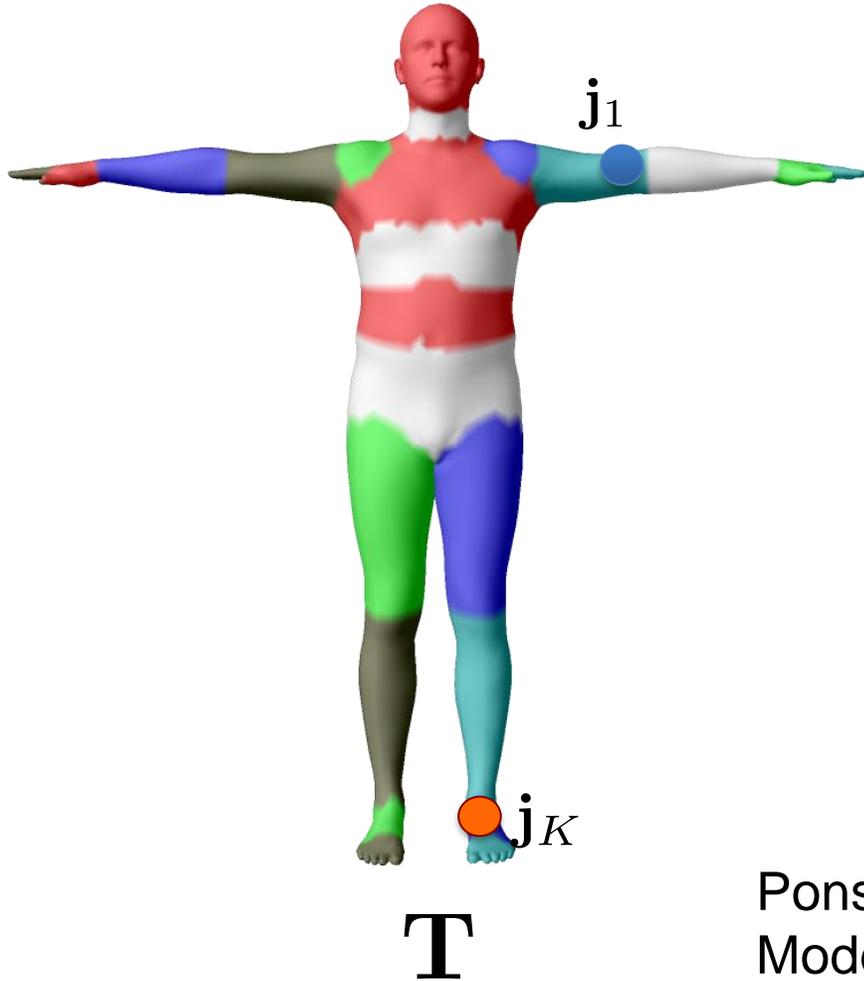
Kinematic Chains



The coordinates of the point in the spatial frame are:

$$\bar{\mathbf{p}}_s = G(\vec{\omega}_1, \vec{\omega}_2, \mathbf{j}_1, \mathbf{j}_2) = G(\vec{\omega}_1, \mathbf{j}_1)G(\vec{\omega}_2, \mathbf{j}_2)\bar{\mathbf{p}}_b$$

Pose Parameters



Given a set of joint locations

$$\mathbf{J} = (\underline{\mathbf{j}}_1, \dots, \underline{\mathbf{j}}_K)^T$$

The pose defined as the vector of concatenated part axis-angles

$$\vec{\theta} = (\underline{\vec{\omega}}_1, \dots, \underline{\vec{\omega}}_k)^T$$

Pons-Moll & Rosenhahn 2011

Model-based Pose Estimation. Looking at People.

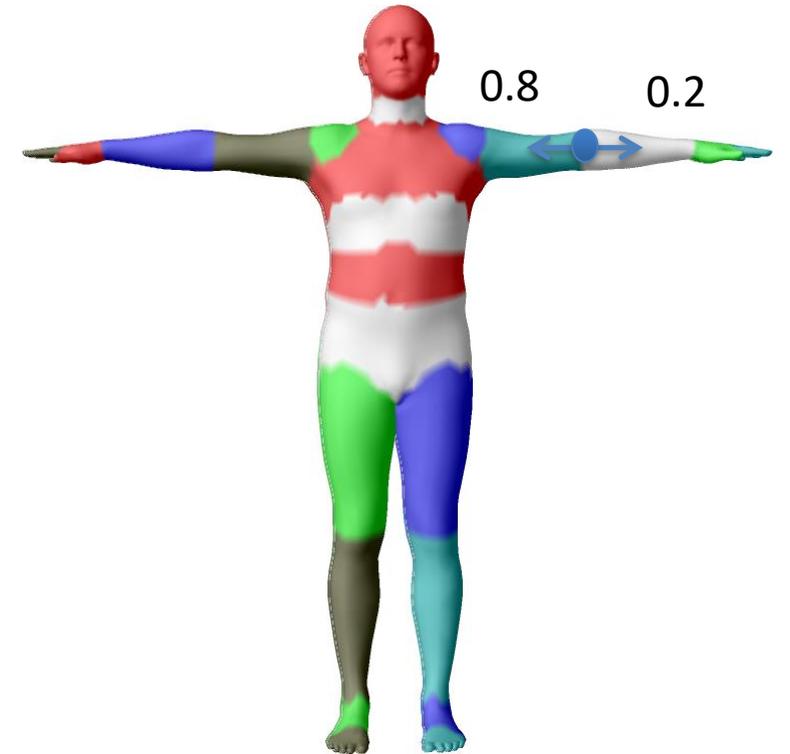
Linear Blend Skinning

$$\bar{\mathbf{t}}'_i = \sum_{k=1}^K w_{k,i} G'_k(\vec{\theta}, \mathbf{J}) \bar{\mathbf{t}}_i$$

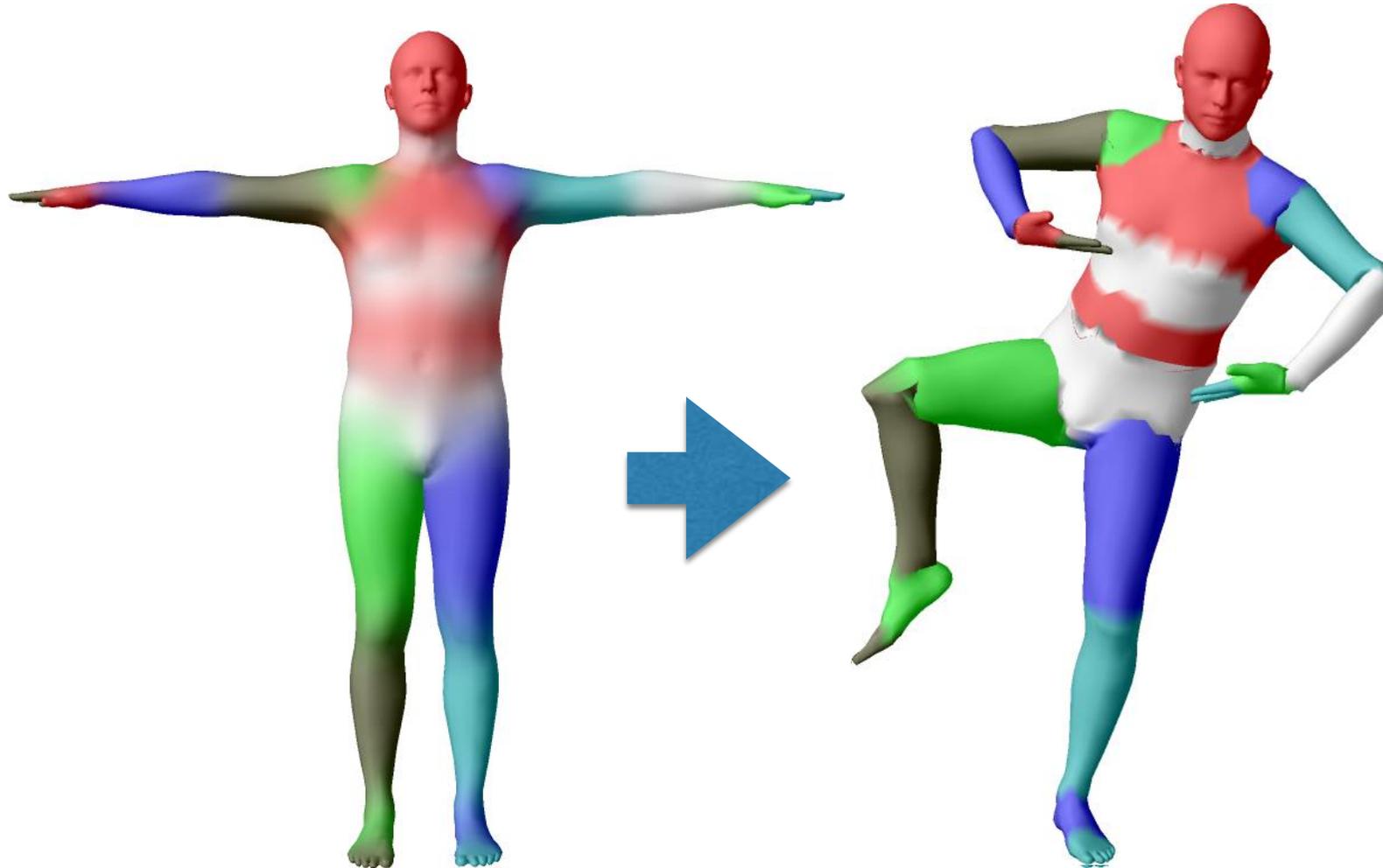
Blend weights

Part transformations

Points transformed as blended linear combination
of joint transformation matrices



Linear Blend Skinning



Standard Skinning

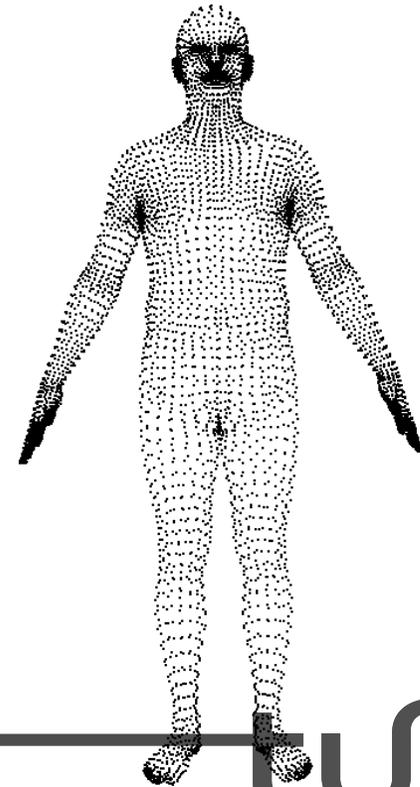
Standard skinning produces vertices from...

– Rest pose vertices: $\mathbf{T} \in \mathbb{R}^{3N}$

– Joint locations: $\mathbf{J} \in \mathbb{R}^{3K}$

– Weights: $\mathcal{W} \in \mathbb{R}^{N \times K}$

– Pose parameters: $\vec{\theta} \in \mathbb{R}^{3K}$



Standard Skinning

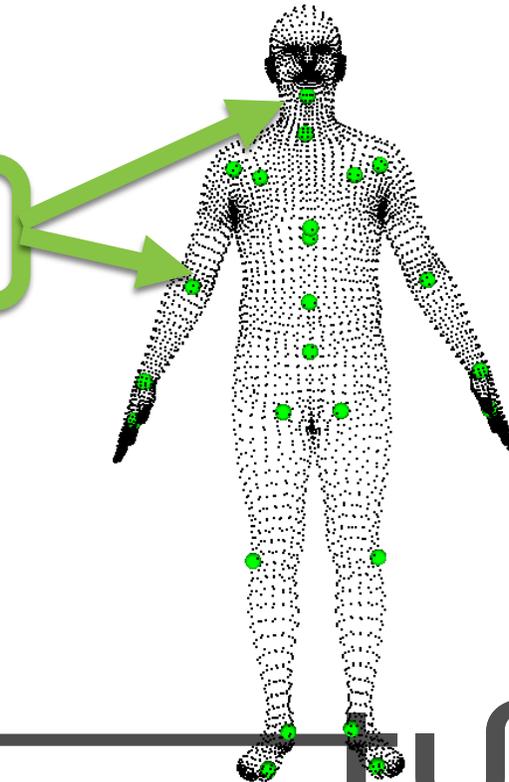
Standard skinning produces vertices from...

– Rest pose vertices: $\mathbf{T} \in \mathbb{R}^{3N}$

– Joint locations: $\mathbf{J} \in \mathbb{R}^{3K}$

– Weights: $\mathcal{W} \in \mathbb{R}^{N \times K}$

– Pose parameters: $\vec{\theta} \in \mathbb{R}^{3K}$



Standard Skinning

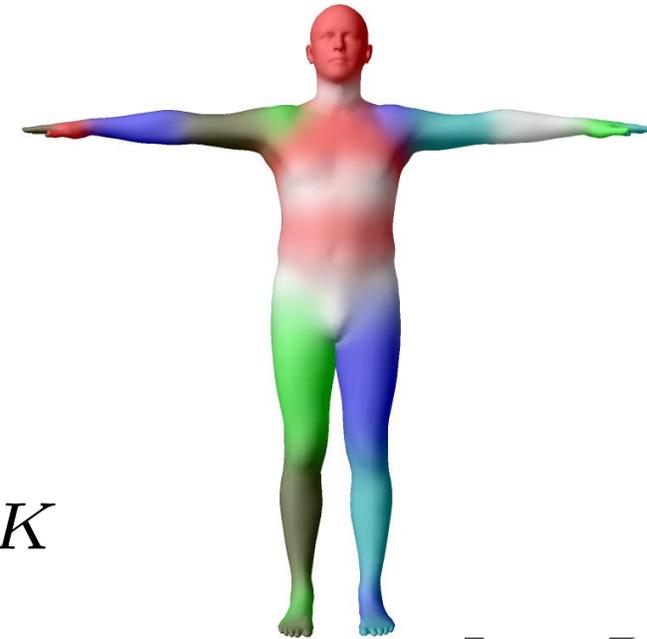
Standard skinning produces vertices from...

– Rest pose vertices: $\mathbf{T} \in \mathbb{R}^{3N}$

– Joint locations: $\mathbf{J} \in \mathbb{R}^{3K}$

– Weights: $\mathcal{W} \in \mathbb{R}^{N \times K}$

– Pose parameters: $\vec{\theta} \in \mathbb{R}^{3K}$



Standard Skinning

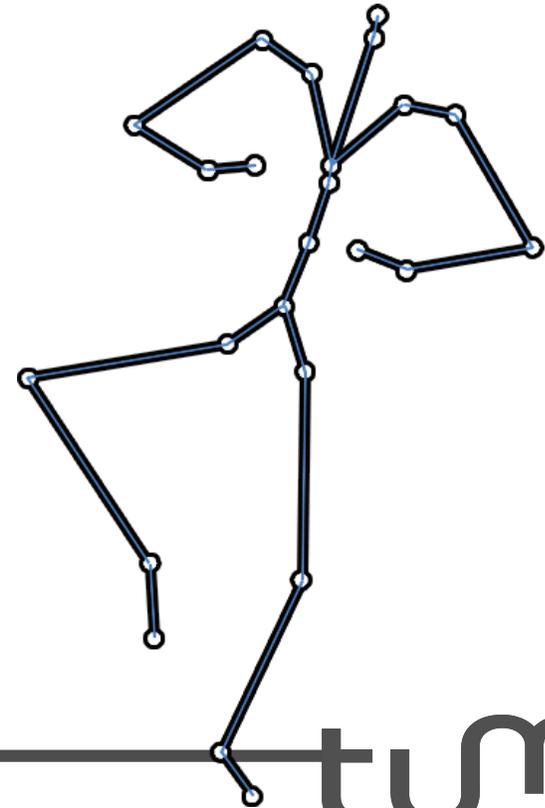
Standard skinning produces vertices from...

– Rest pose vertices: $\mathbf{T} \in \mathbb{R}^{3N}$

– Joint locations: $\mathbf{J} \in \mathbb{R}^{3K}$

– Weights: $\mathcal{W} \in \mathbb{R}^{N \times K}$

– Pose parameters: $\vec{\theta} \in \mathbb{R}^{3K}$

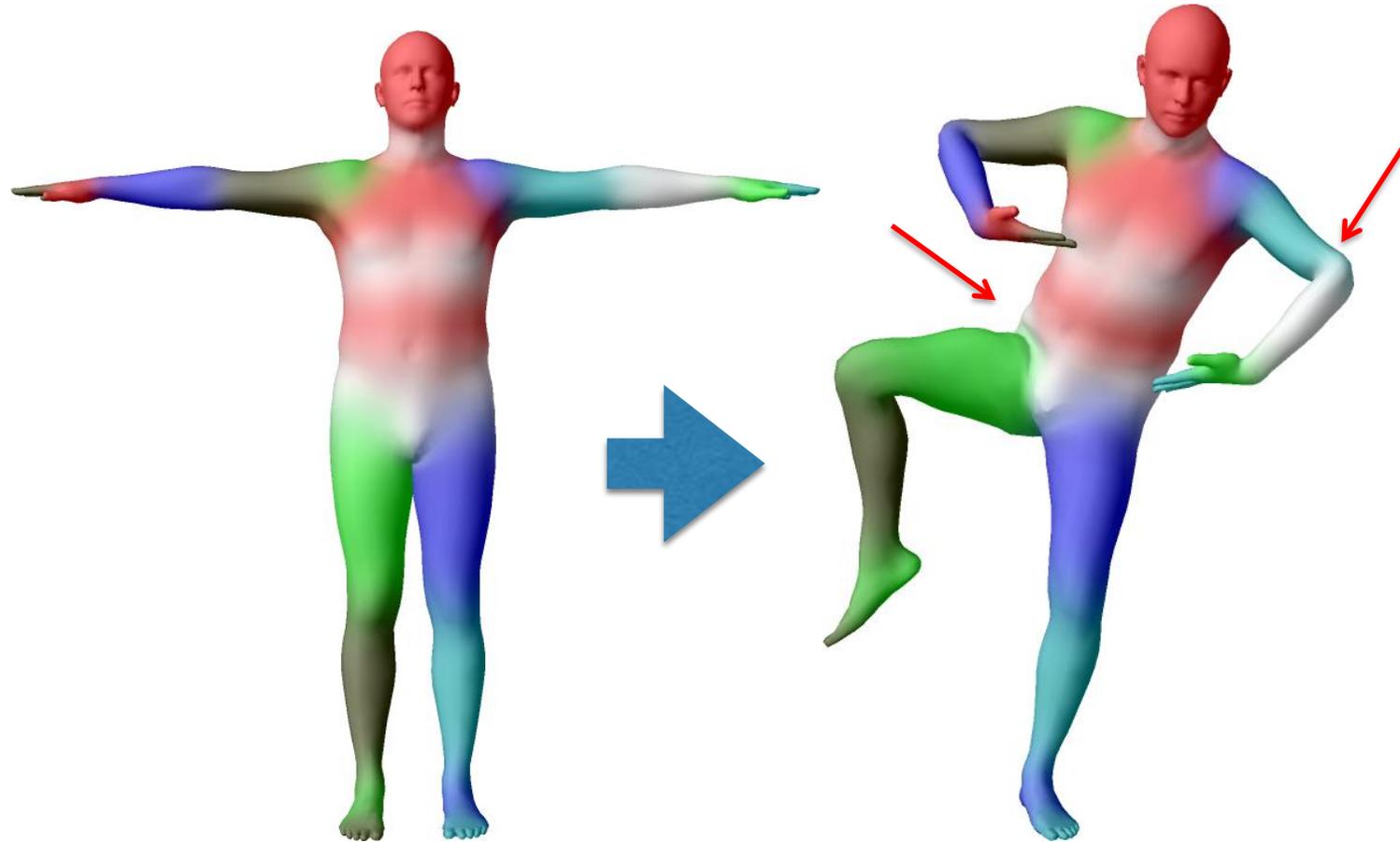


Skinning Function

- Rest pose vertices: $\mathbf{T} \in \mathbb{R}^{3N}$
- Joint locations: $\mathbf{J} \in \mathbb{R}^{3K}$
- Weights: $\mathcal{W} \in \mathbb{R}^{N \times K}$
- Pose parameters: $\vec{\theta} \in \mathbb{R}^{3K}$

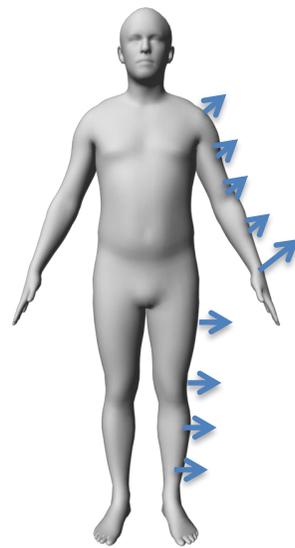
$$W(\mathbf{T}, \mathbf{J}, \mathcal{W}, \vec{\theta}) \mapsto \text{vertices}$$

LBS Problems



Solution: (Corrective) Blend Shapes

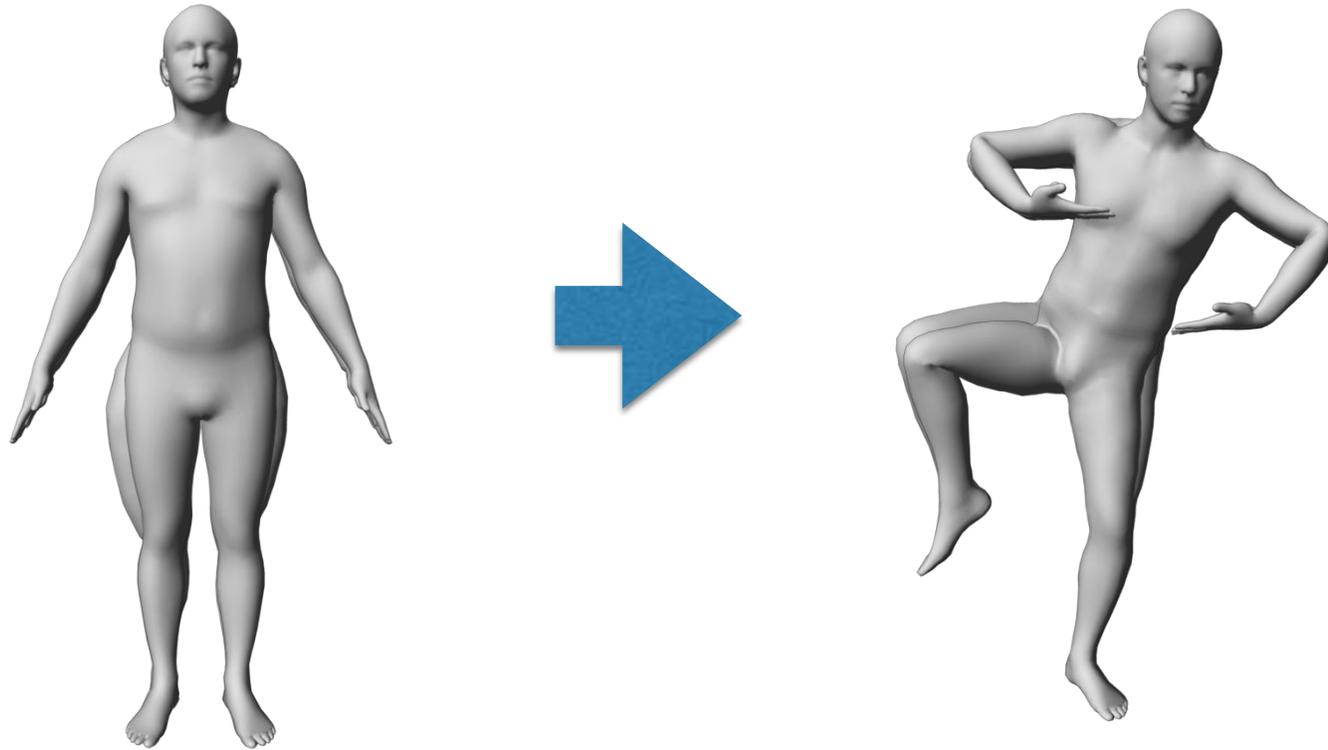
- A **blend shape** is a set of vertex displacements in a rest pose
 - Pose blend shapes: correct for LBS problems



$$\mathbf{P} = \text{vec} \left(\begin{bmatrix} \Delta x_1 & \Delta y_1 & \Delta z_1 \\ \vdots & \vdots & \vdots \\ \Delta x_N & \Delta y_N & \Delta z_N \end{bmatrix} \right) \rightarrow \text{Offset 1} \in \mathbb{R}^{3N}$$

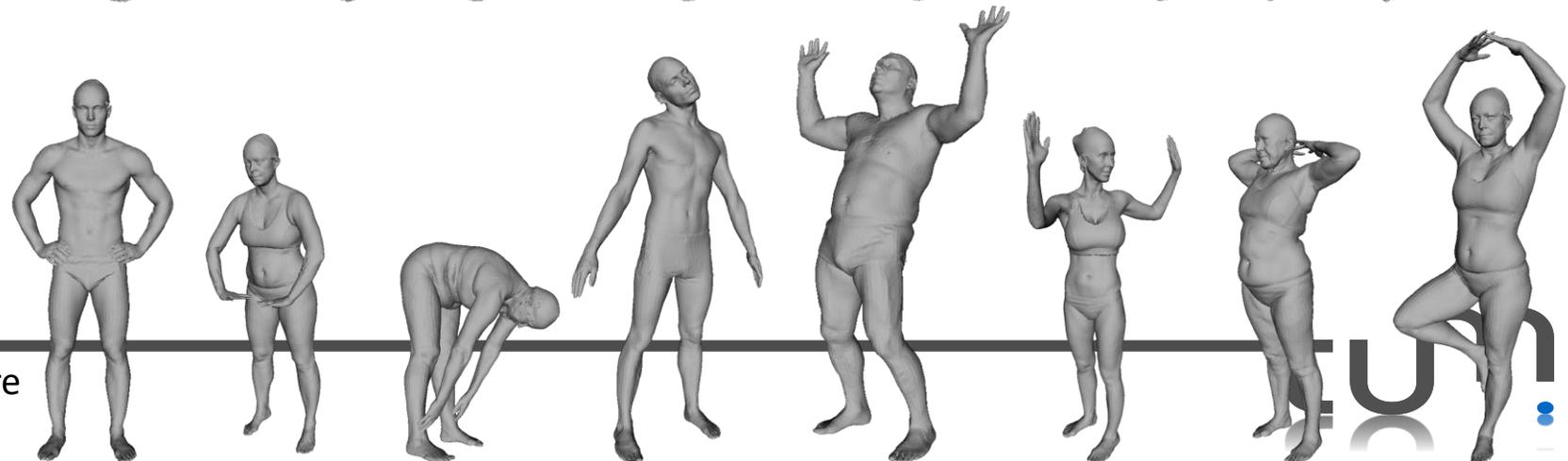
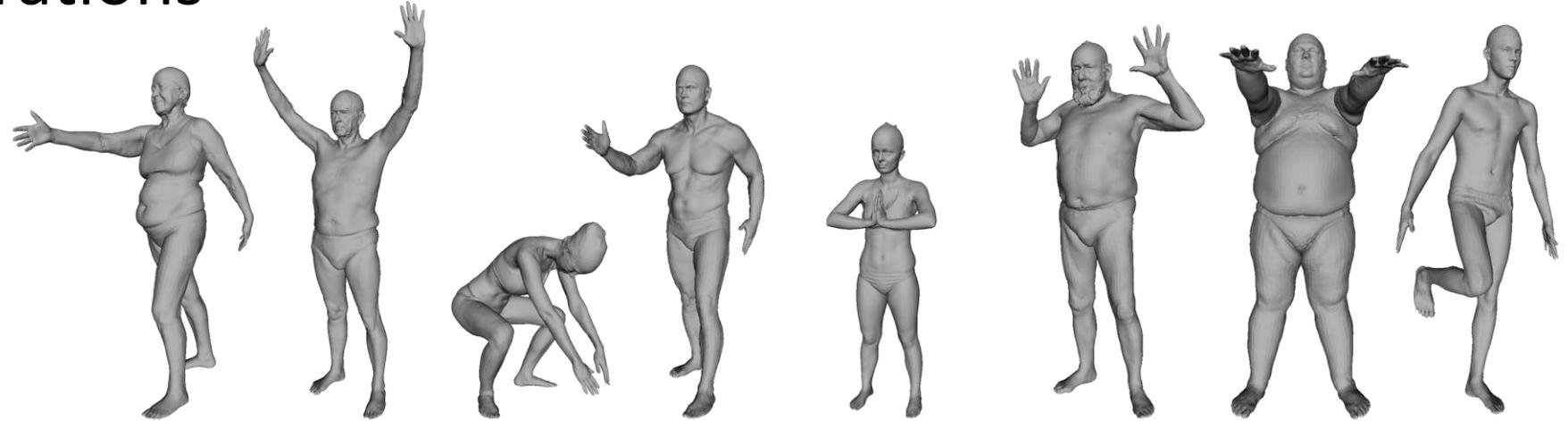
Pose Blend Shapes

- **With** blend shape correction



SMPL: Skinned Multi-Person Linear Model

- Multipose database: 20 males, 24 females
- 1800 registrations



[Loper et al. 15]

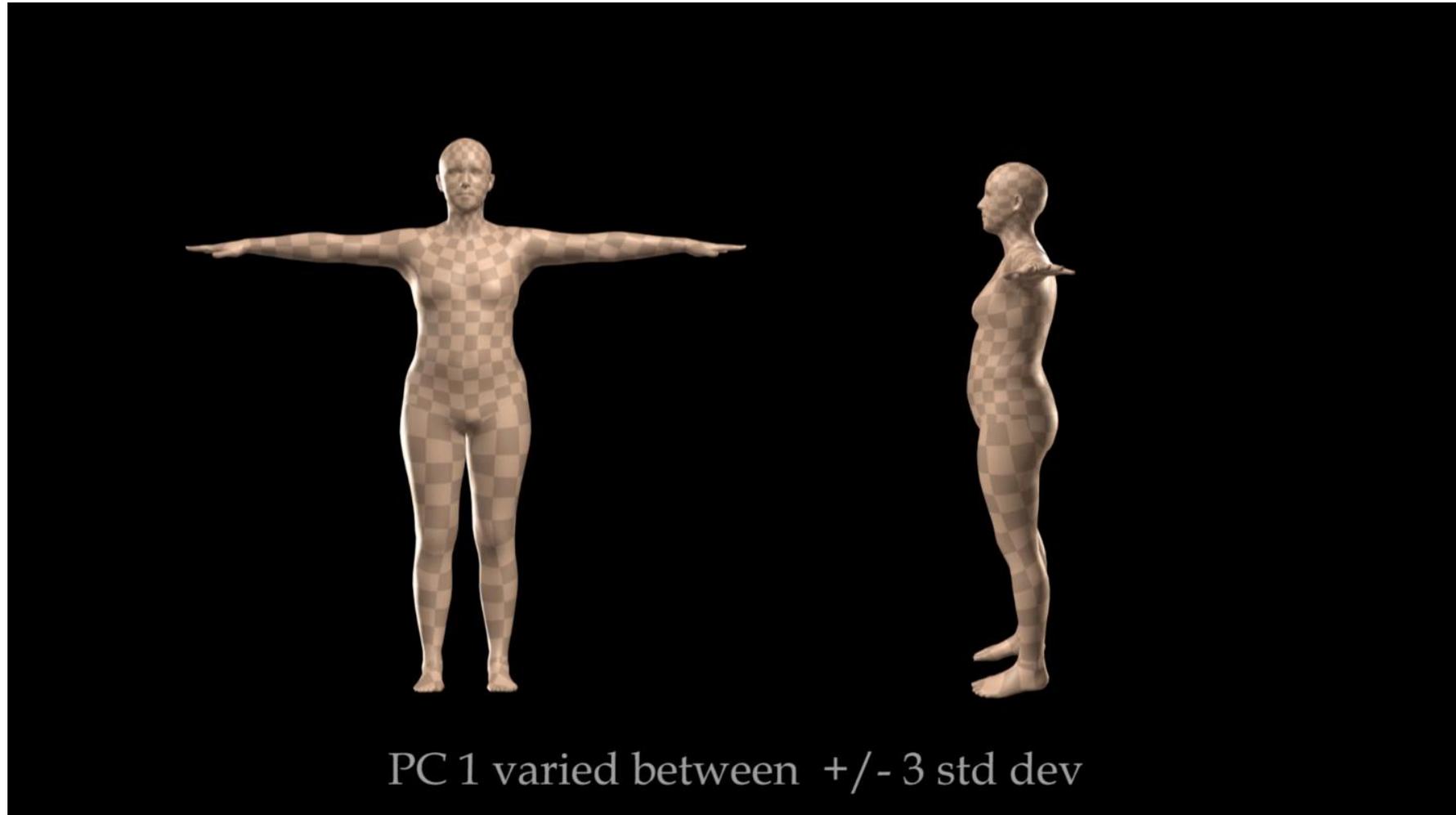
SMPL: Skinned Multi-Person Linear Model

- Multishape database: PCA on ~ 2000 single-pose registrations per gender

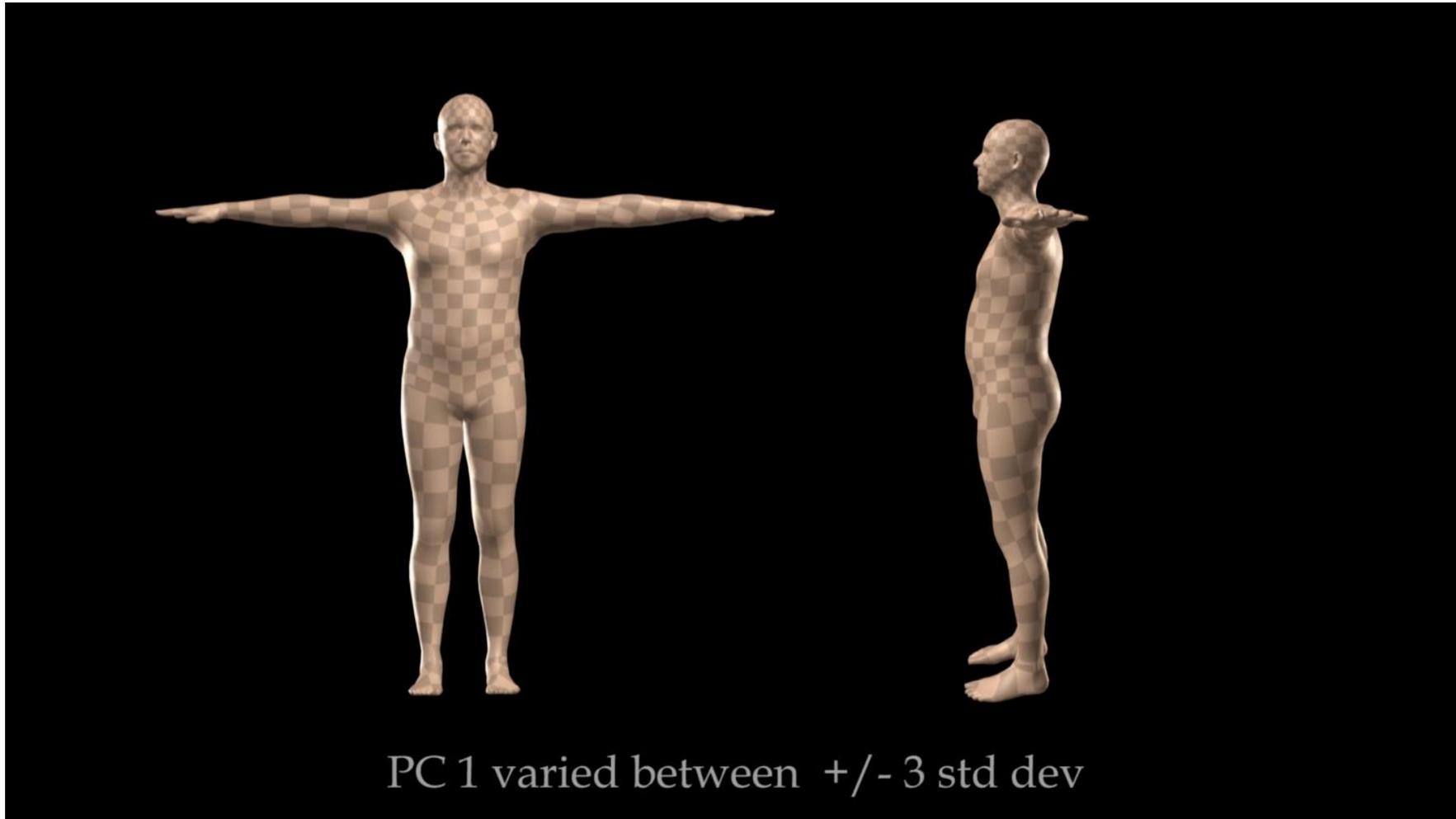


Multi-shape Training Set

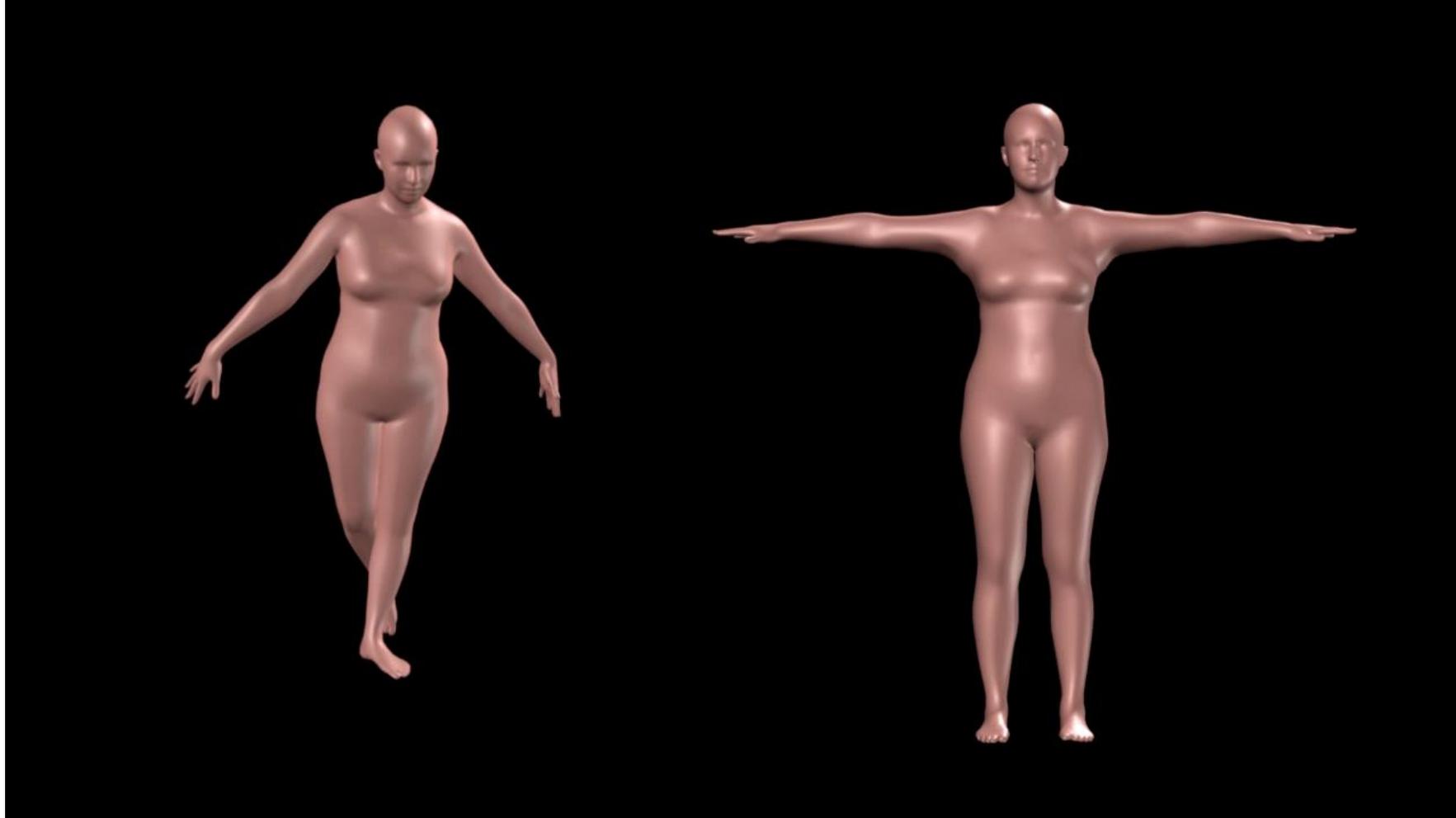
Shape Blend Shapes- Female



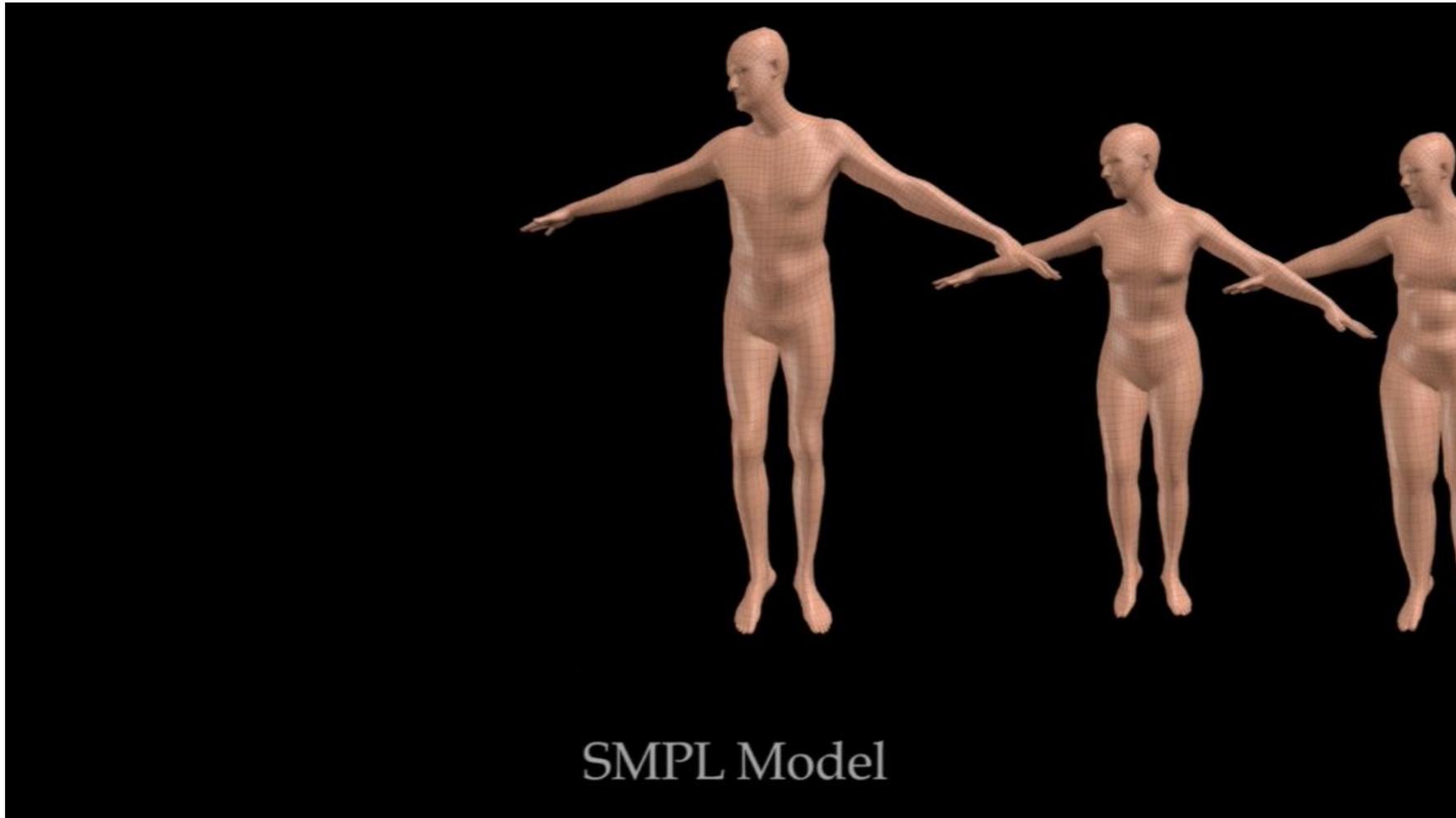
Shape Blend Shapes- Male



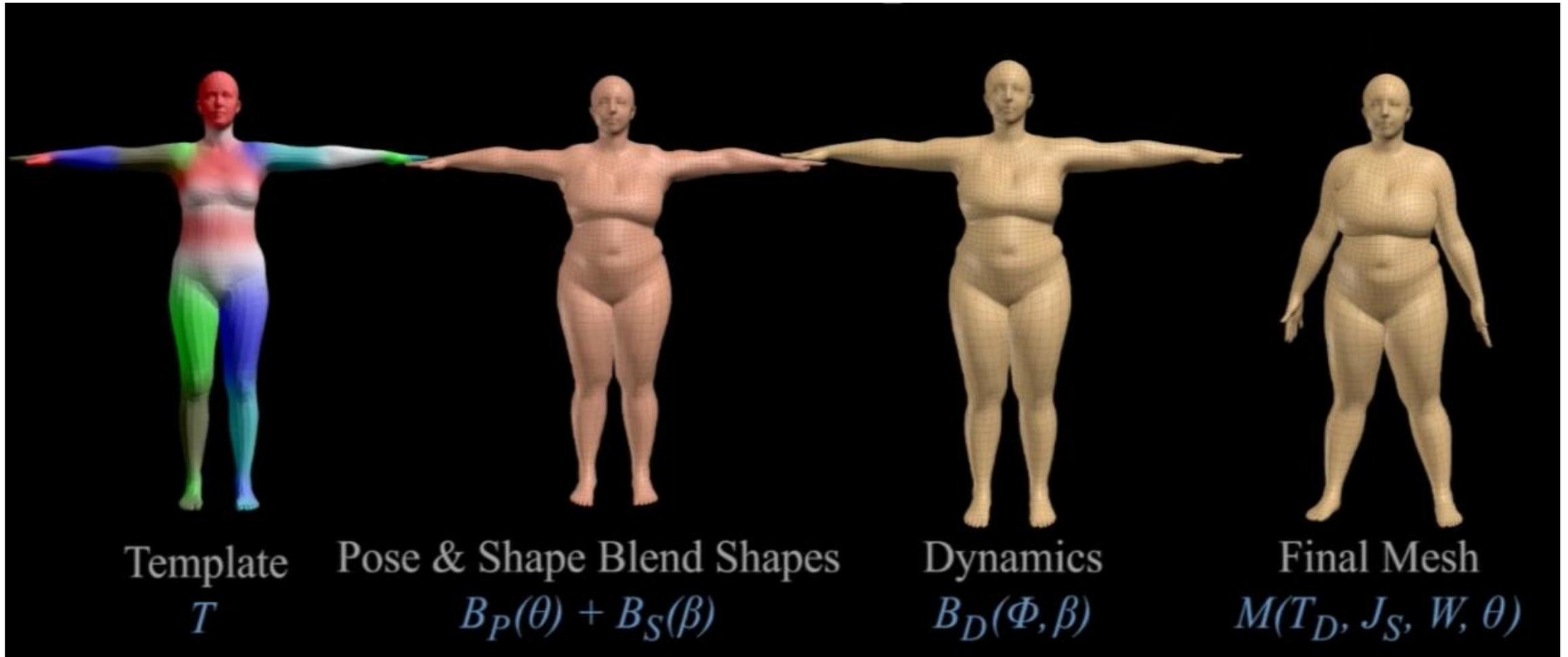
Pose Blendshapes



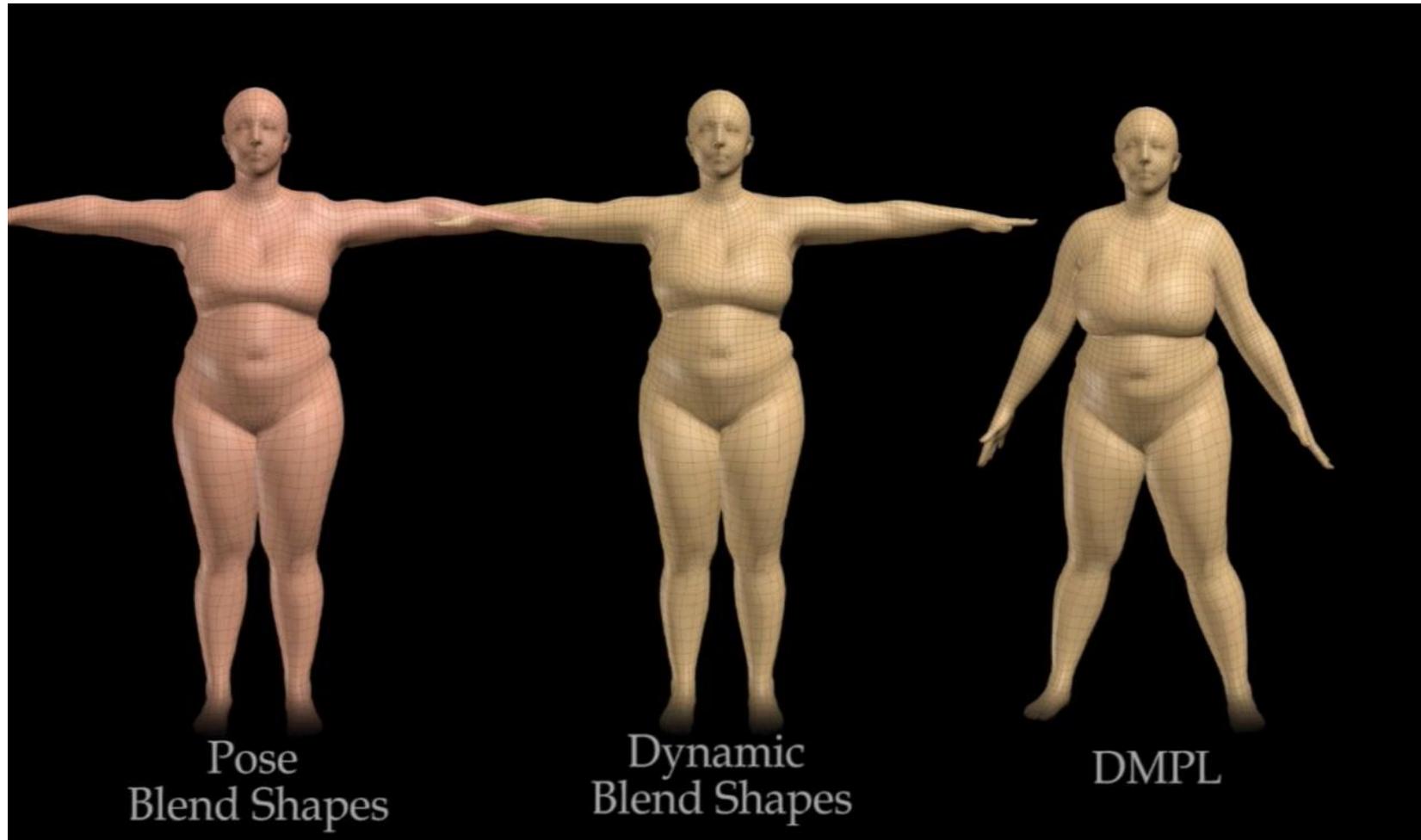
SMPL Results



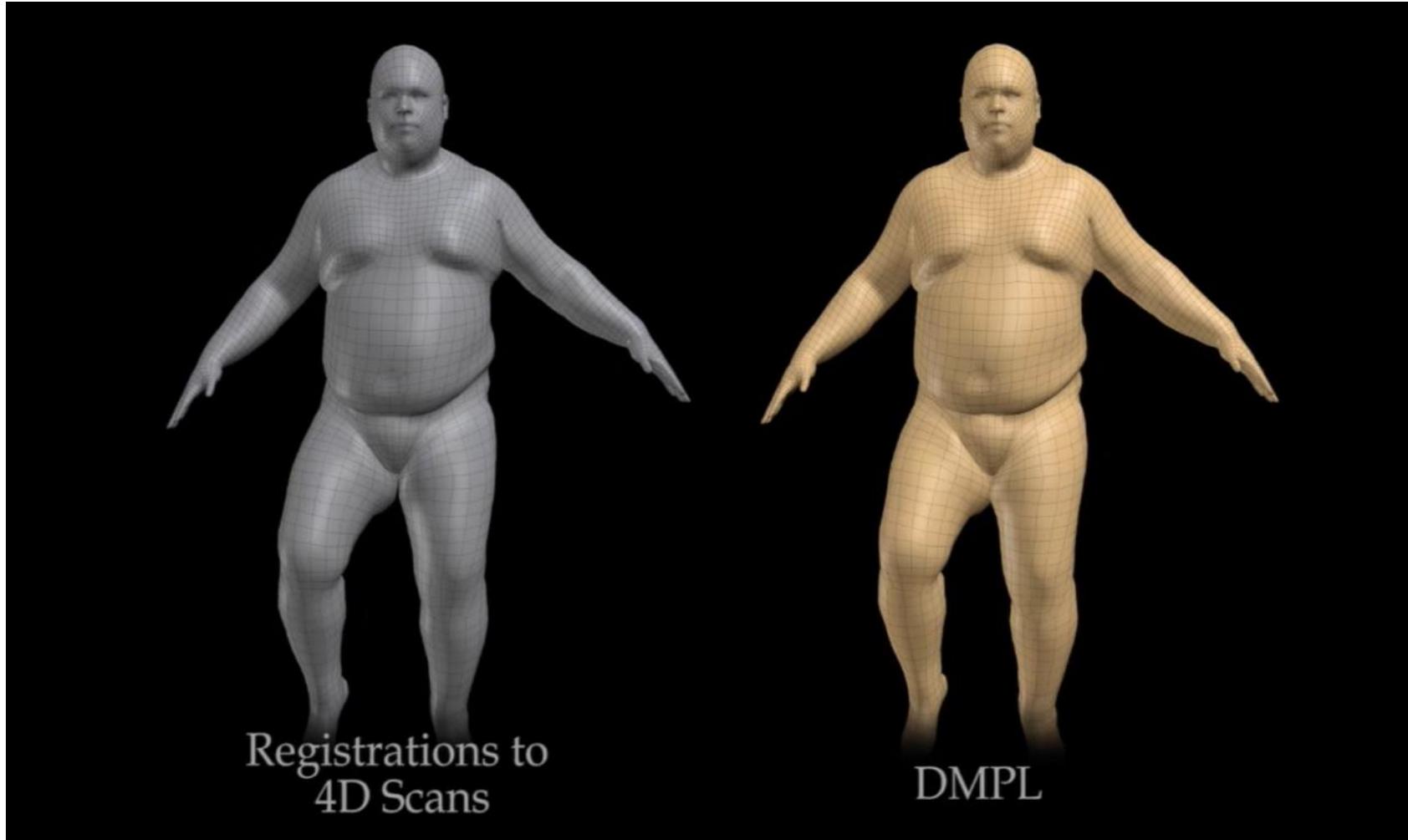
Model Decomposition



Dynamics of Soft Tissue



DMPL Exaggeration



Reconstruction of Human Bodies

- Non-rigid registration to align scans

– Fit template to input scan: $E = w_d E_d + w_s E_s + w_m E_m$

Data error: template and target surfaces as close as possible

$$E_d = \sum_{i=1}^n w_i \text{dist}^2(\mathbf{T}_i \mathbf{v}_i, D)$$

Transformation smoothness

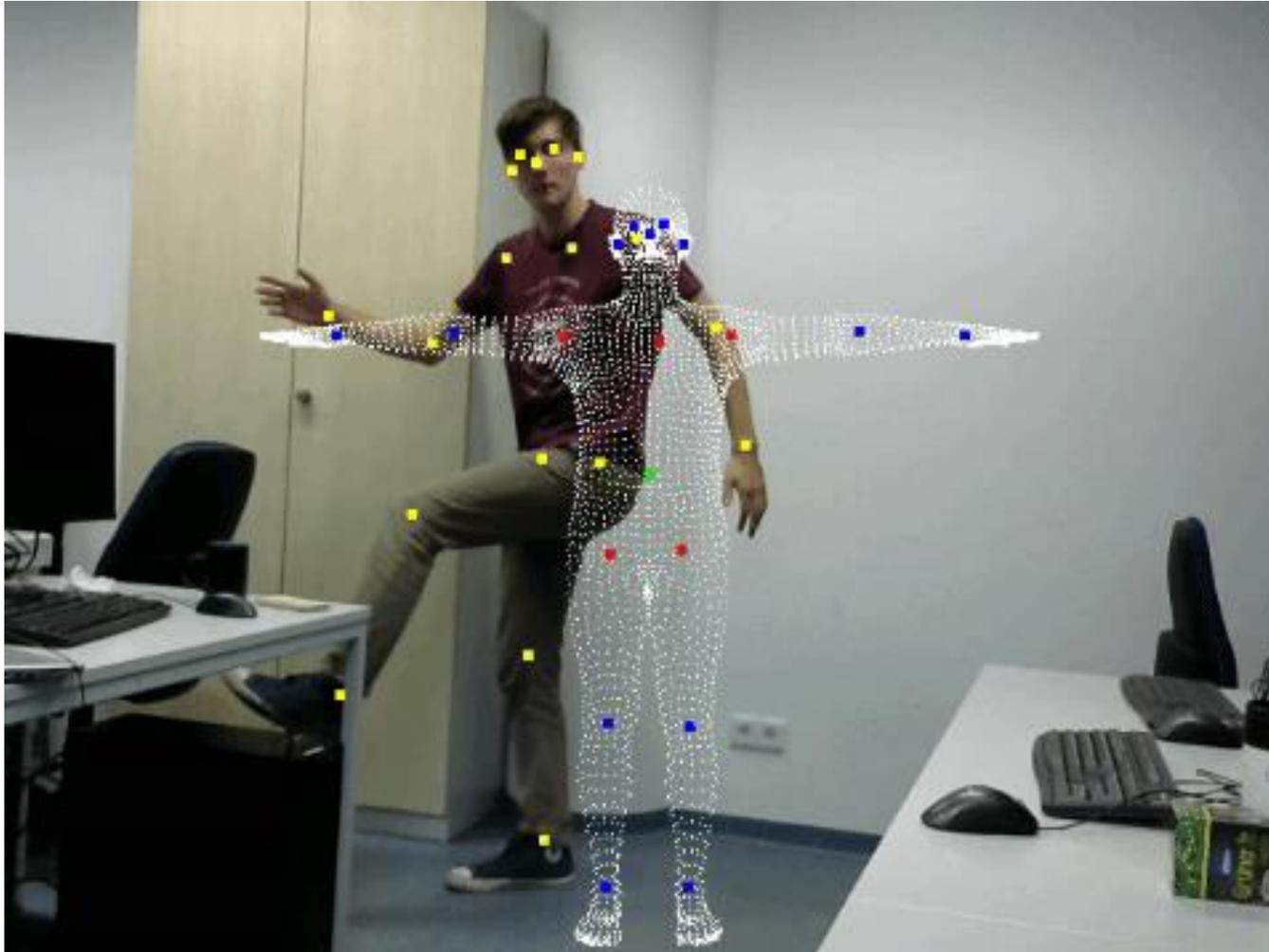
$$E_s = \sum_{\{i,j\} | \{\mathbf{v}_i, \mathbf{v}_j\} \in \text{edges}(\mathcal{T})} \|\mathbf{T}_i - \mathbf{T}_j\|_F^2$$

Marker error

$$E_m = \sum_{i=1}^m \|\mathbf{T}_{\kappa_i} \mathbf{v}_{\kappa_i} - \mathbf{m}_i\|^2$$



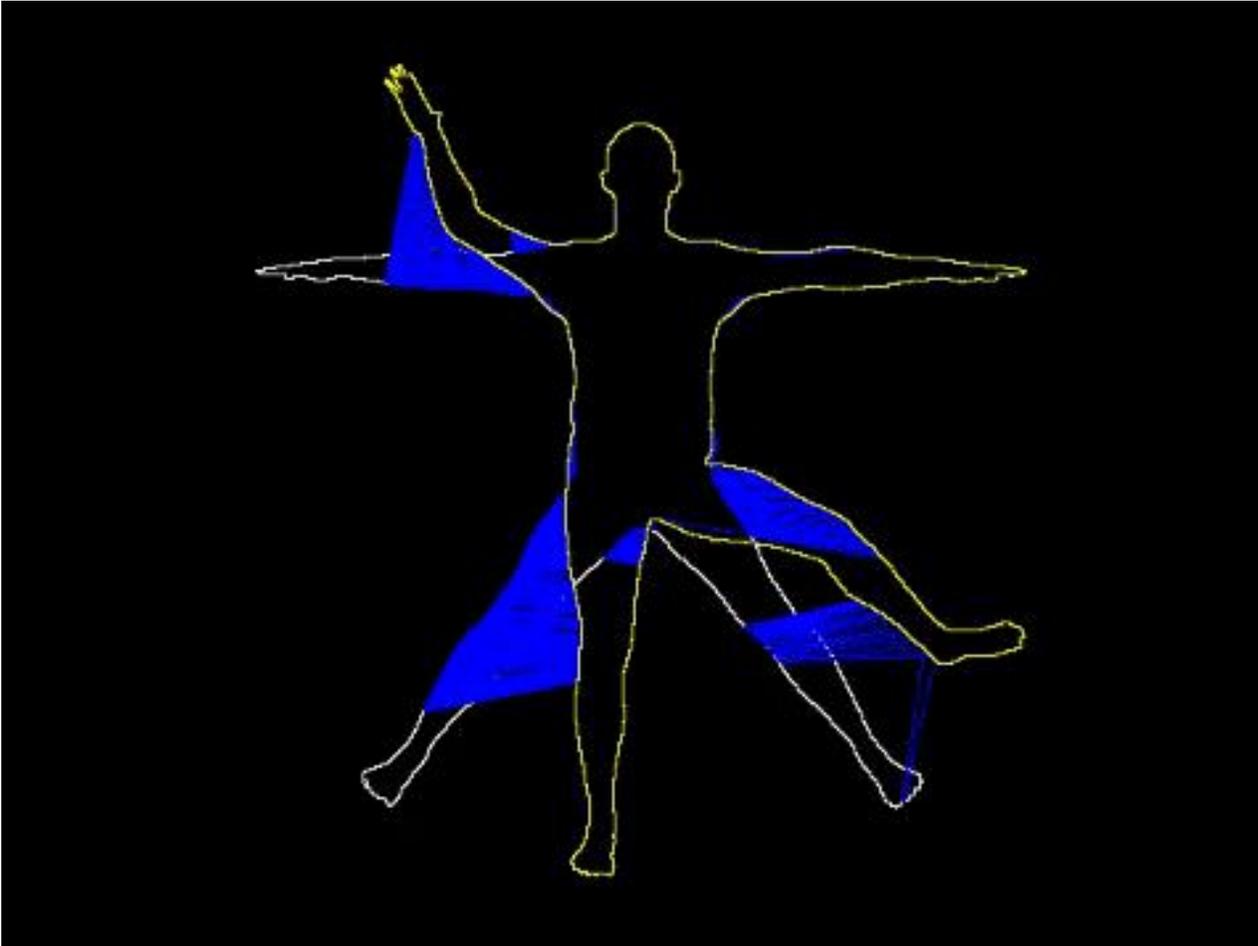
Reconstruction of Human Bodies



Marker error

$$E_m = \sum_{i=1}^m \|\mathbf{T}_{\kappa_i} \mathbf{v}_{\kappa_i} - \mathbf{m}_i\|^2$$

Reconstruction of Human Bodies



Data error: template and target surfaces as close as possible

$$E_d = \sum_{i=1}^n w_i \text{dist}^2(\mathbf{T}_i \mathbf{v}_i, \mathcal{D})$$

Here: Silhouette constraints

Reconstruction of Human Bodies



Video Based Reconstruction of 3D People Models

Thiemo Alldieck¹, Marcus Magnor¹, Weipeng Xu², Christian Theobalt², Gerard Pons-Moll²

¹ Computer Graphics Lab, TU Braunschweig

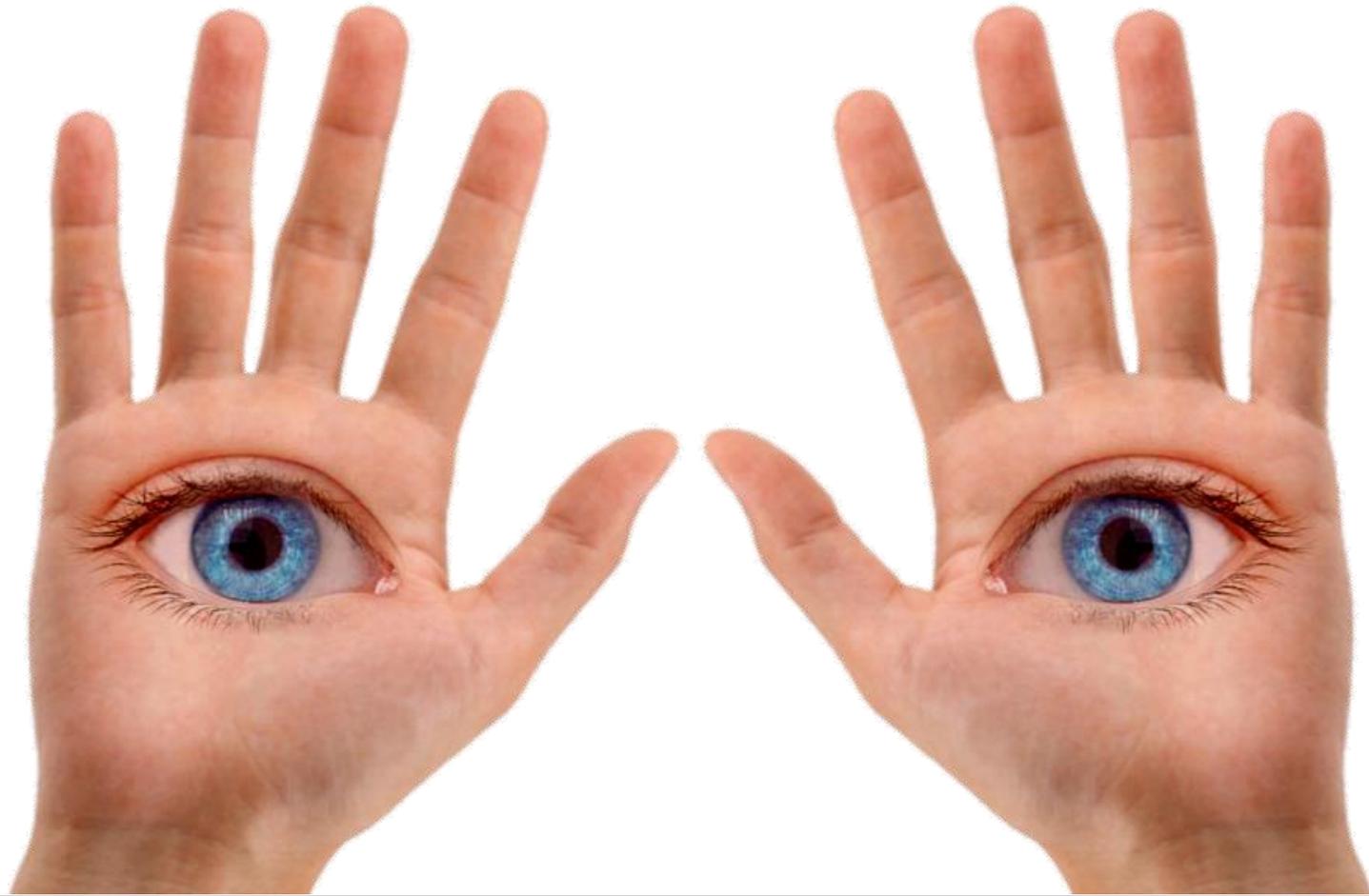
² Max Planck Institute for Informatics



Computer Graphics
TU Braunschweig

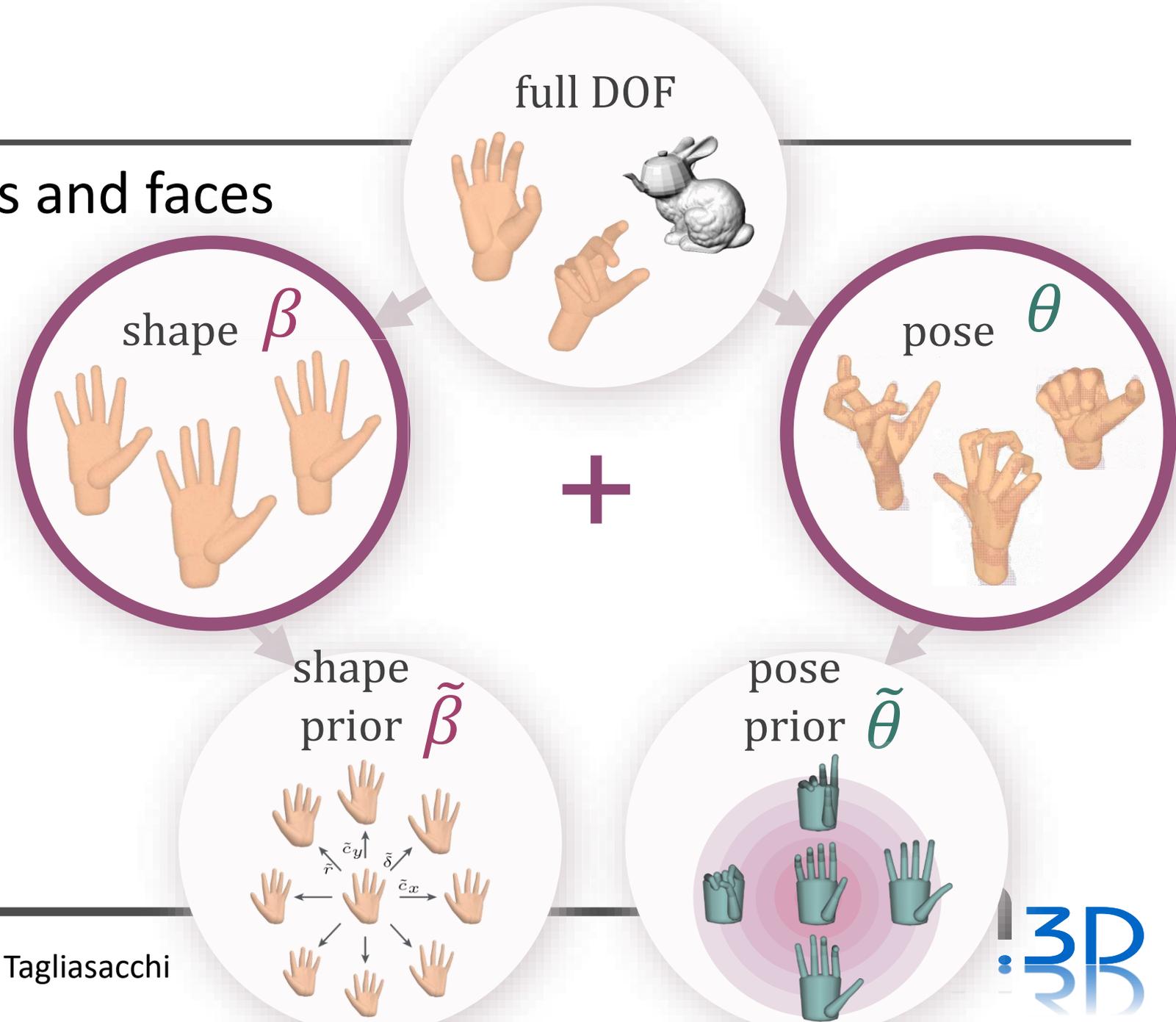


Hands

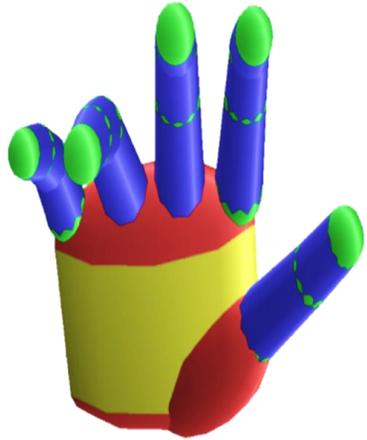


Hand Tracking

- Idea similar to bodies and faces
 - Statistical models
 - Anchored in skel.
 - Training data
 - Fitting (lower dim)



Hand Models



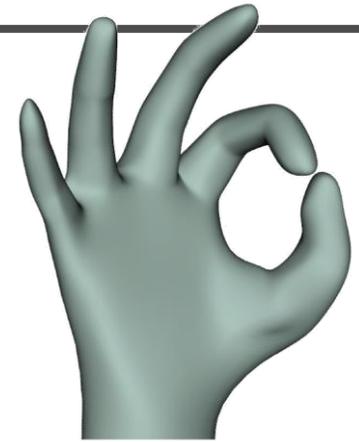
Oikonomidis et al. 2011



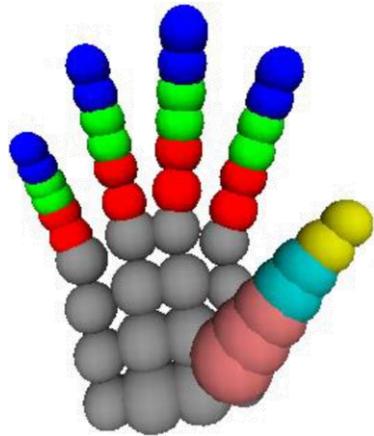
Melax et al. 2013



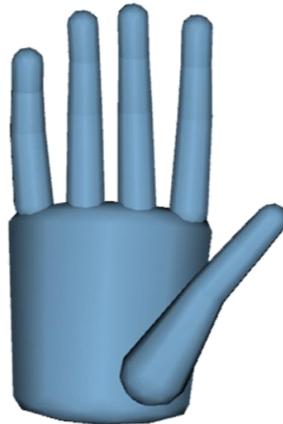
Sharp et al. 2015



Taylor et al. 2016



Qian et al. 2014



Tagliasacchi et al. 2015



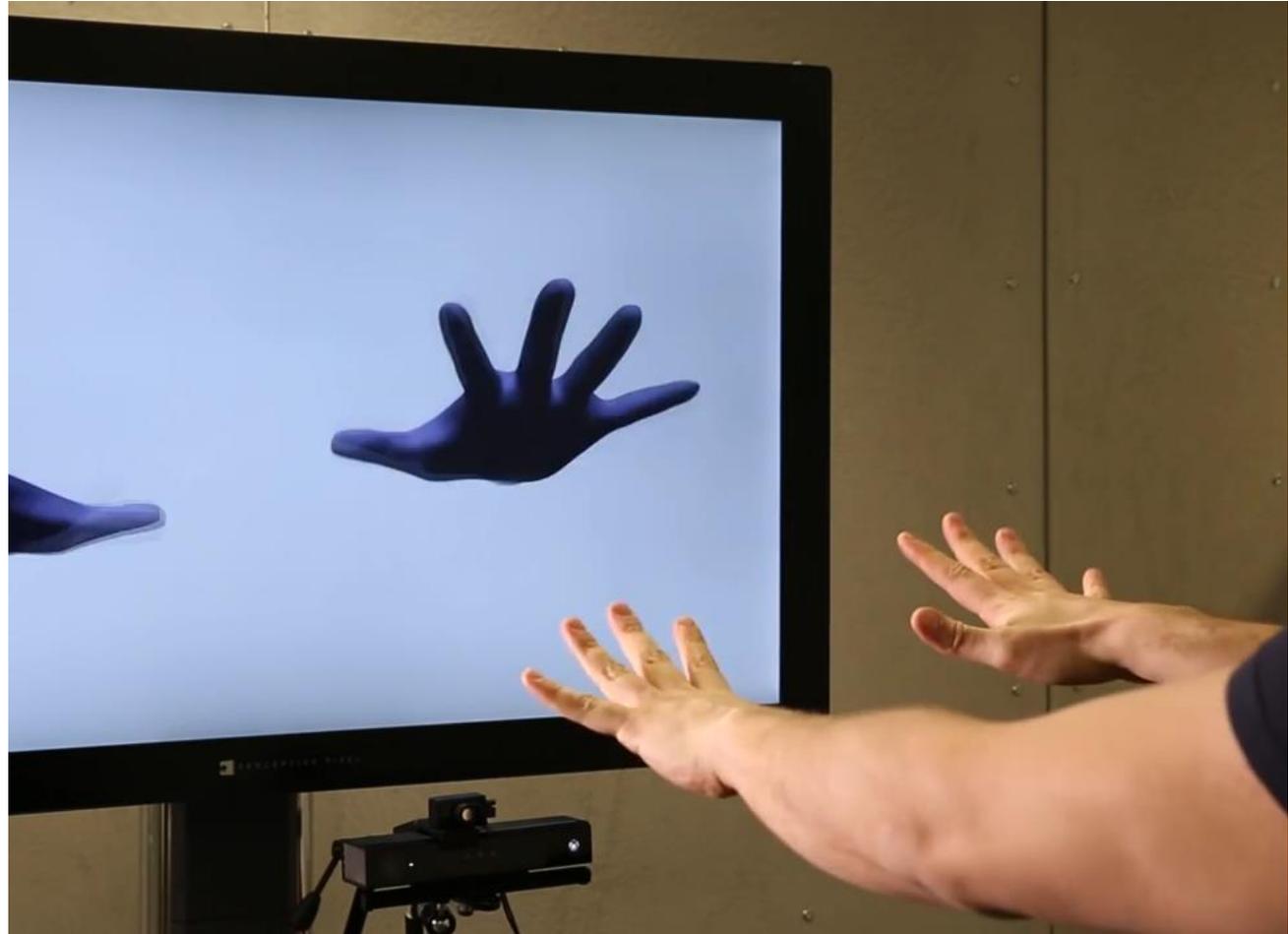
Khamis et al. 2015



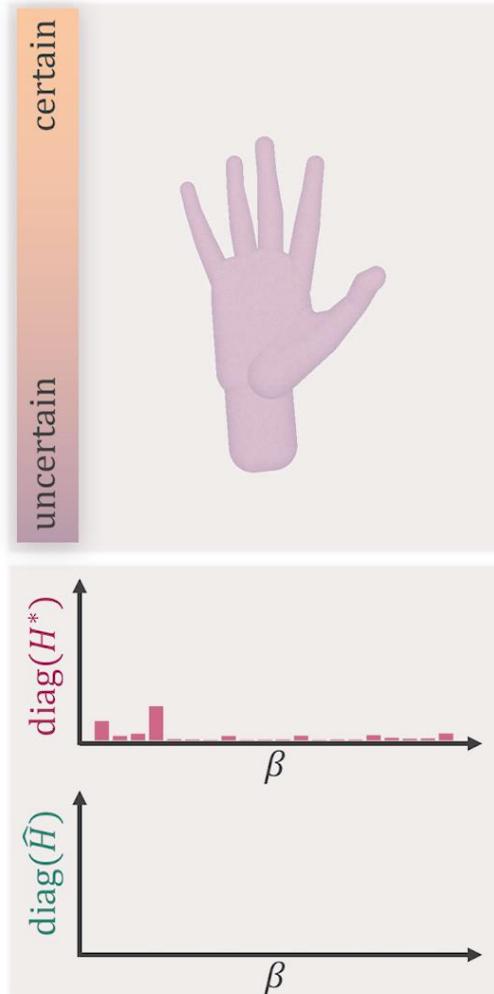
Tkach et al. 2016

[Taylor et al. 16]

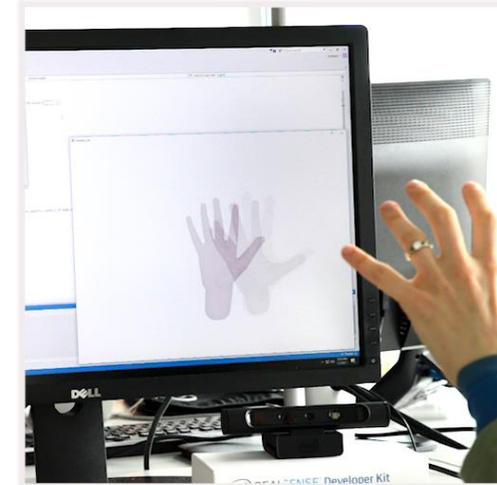
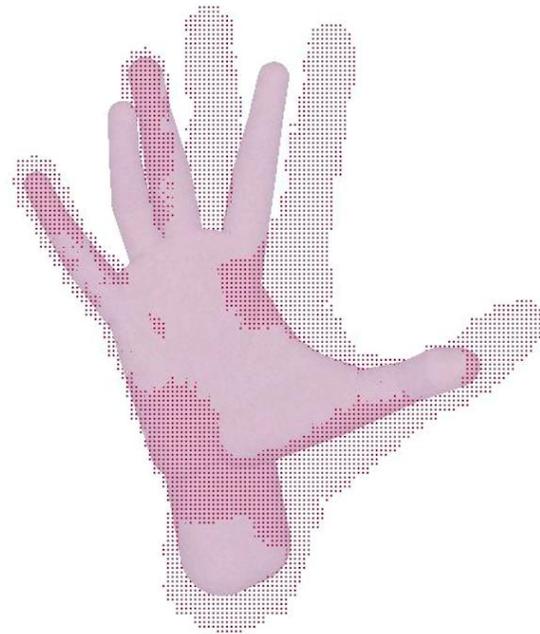
- Fitting using a subdiv model
 - C2 continuous (mostly)
 - Allows super fast LM fit!



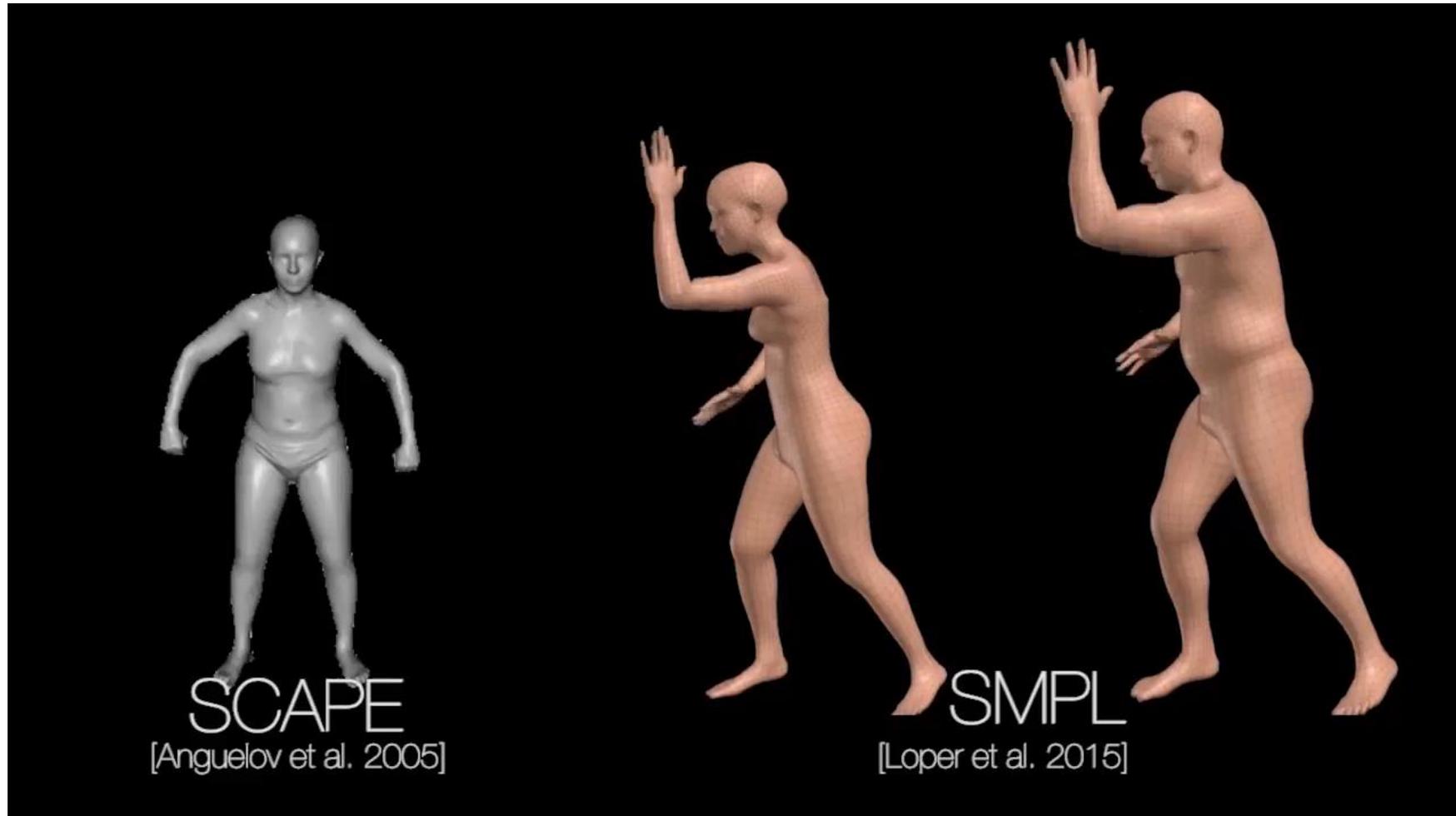
Online Calibration



Online Calibration



Body + Hands



Model Generation and Fitting Issues

- Need training data -> i.e., expensive scanning setup
- Neutral pose between subjects is similar but not identical
 - i.e., a 'zero'-pose is not always the same
- Ambiguities for fitting: identity or pose change?
 - Can fit input by changing the body's shape *or* the expression
- Drift in Non-rigid Registration (during model generation)

Some Cool Links (Body Tracking -- incomplete)

- CESAR DB:
 - <http://store.sae.org/caesar/>
- SCAPE [Anguelov et al. 03]
 - <http://ai.stanford.edu/~drago/Projects/scape/scape.html>
- SMPL: A Skinned Multi-Person Linear Model [Loper et al. 15]
 - <https://www.youtube.com/watch?v=kuBlUyHeV5U>

-
- PIFu:
 - <https://shunsukesaito.github.io/PIFu/>
 - NASA:
 - <http://virtualhumans.mpi-inf.mpg.de/papers/NASA20/NASA.pdf>

Some Links (Hand Tracking -- incomplete)

- Really cool tutorial on fitting surface to data (from an optimization perspective) [Fitzgibbon and Taylor 16]
 - http://awf.fitzgibbon.ie/cvpr16_tutorial
- Accurate, Robust, and Flexible Real-time Tracking [Sharp et al. 15]
 - <https://www.youtube.com/watch?v=RQ-kAoaNc60>
- Efficient and Precise Interactive Hand Tracking [Taylor et al. 16]
 - <https://www.youtube.com/watch?v=QTz1zQAnMcU>
- Sphere Meshes for Real-time Hand Modeling and Tracking [Tkach et al. 16]
 - <https://www.youtube.com/watch?v=QtOQmbo3IsY>
- Online Generative Model Personalization for Hand Tracking [Tkach et al. 17]
 - <https://www.youtube.com/watch?v=zbcoWcYg4Qs>