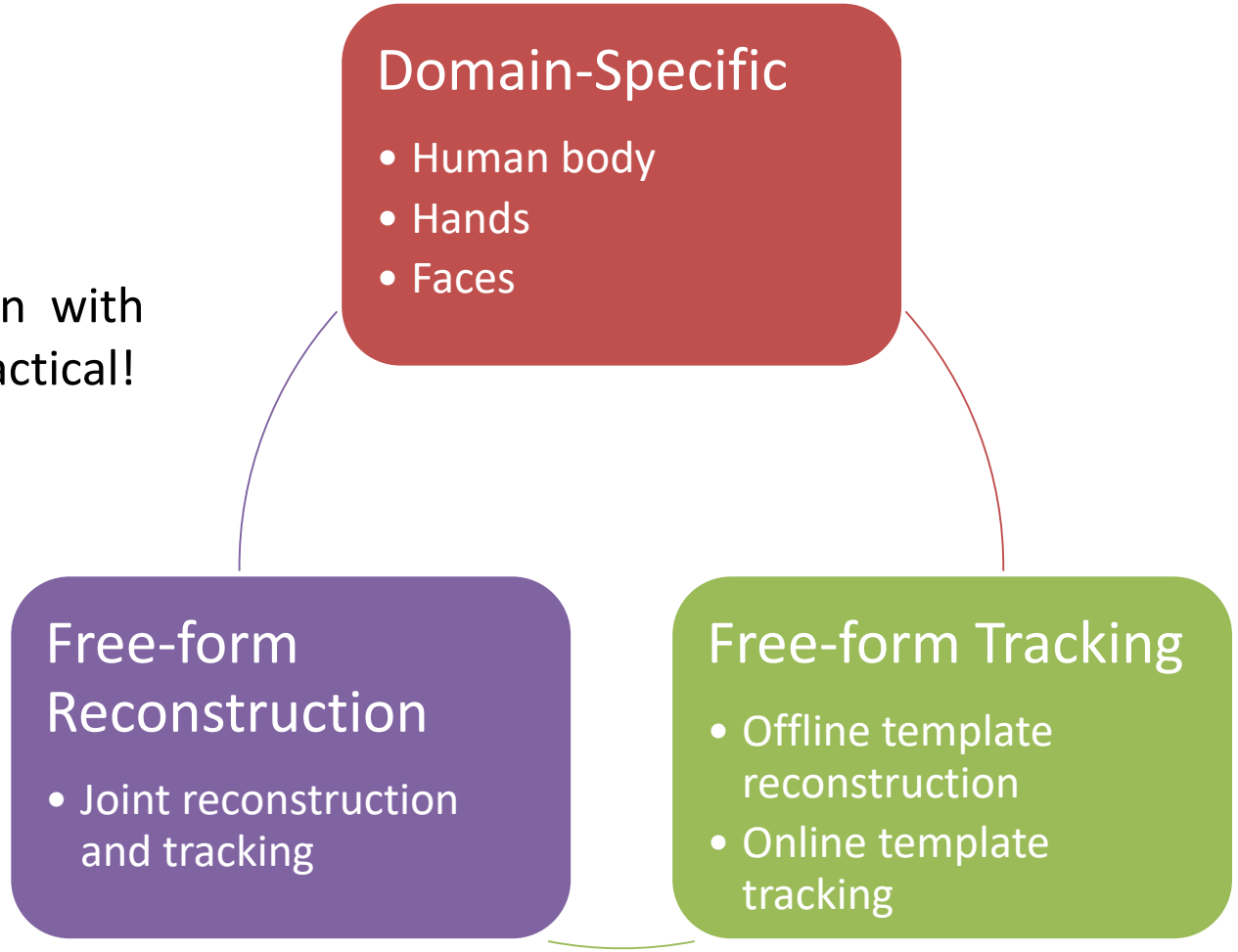
3D Scanning & Motion Capture

Body and Hand Tracking

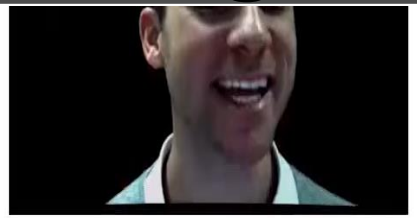
Dr. Justus Thies, Dr. Angela Dai

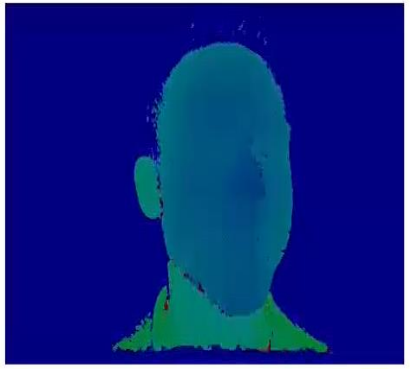
Dynamic 3D Capture

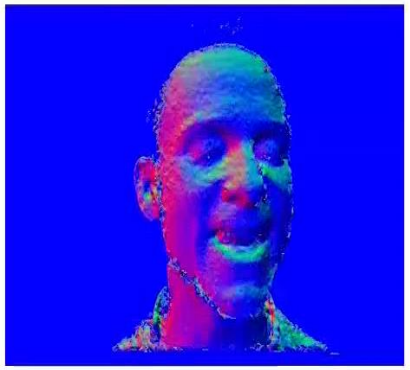
Need more regularization with fewer DoF to make it practical!



Non-rigid Tracking

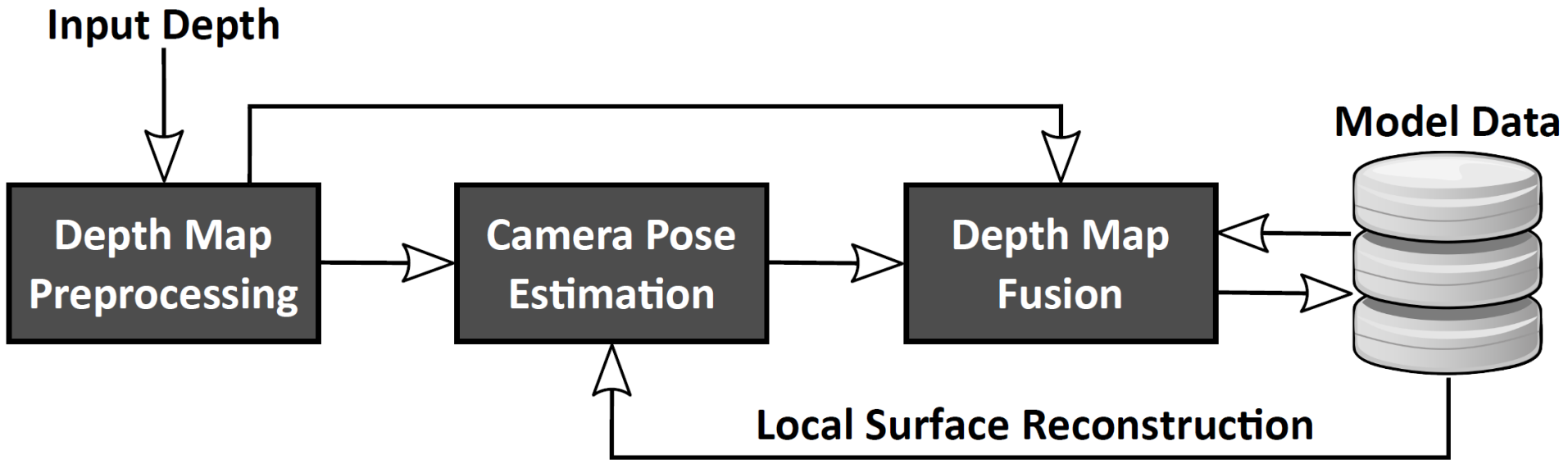


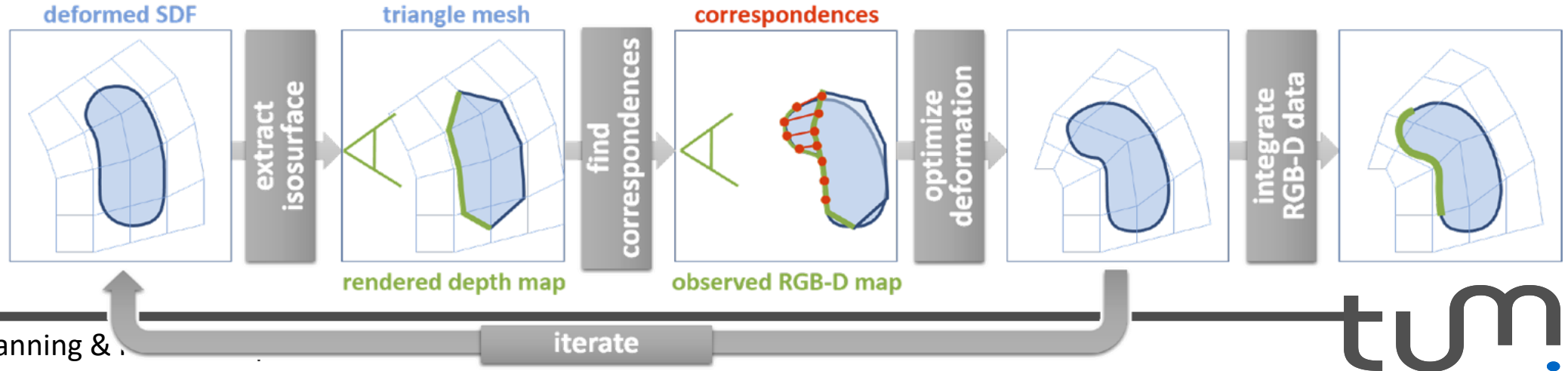




TOG'14 [Zollhöfer et al.]: Deformables

Rigid vs Non-Rigid Reconstruction





Non-rigid Reconstruction

DynamicFusion:

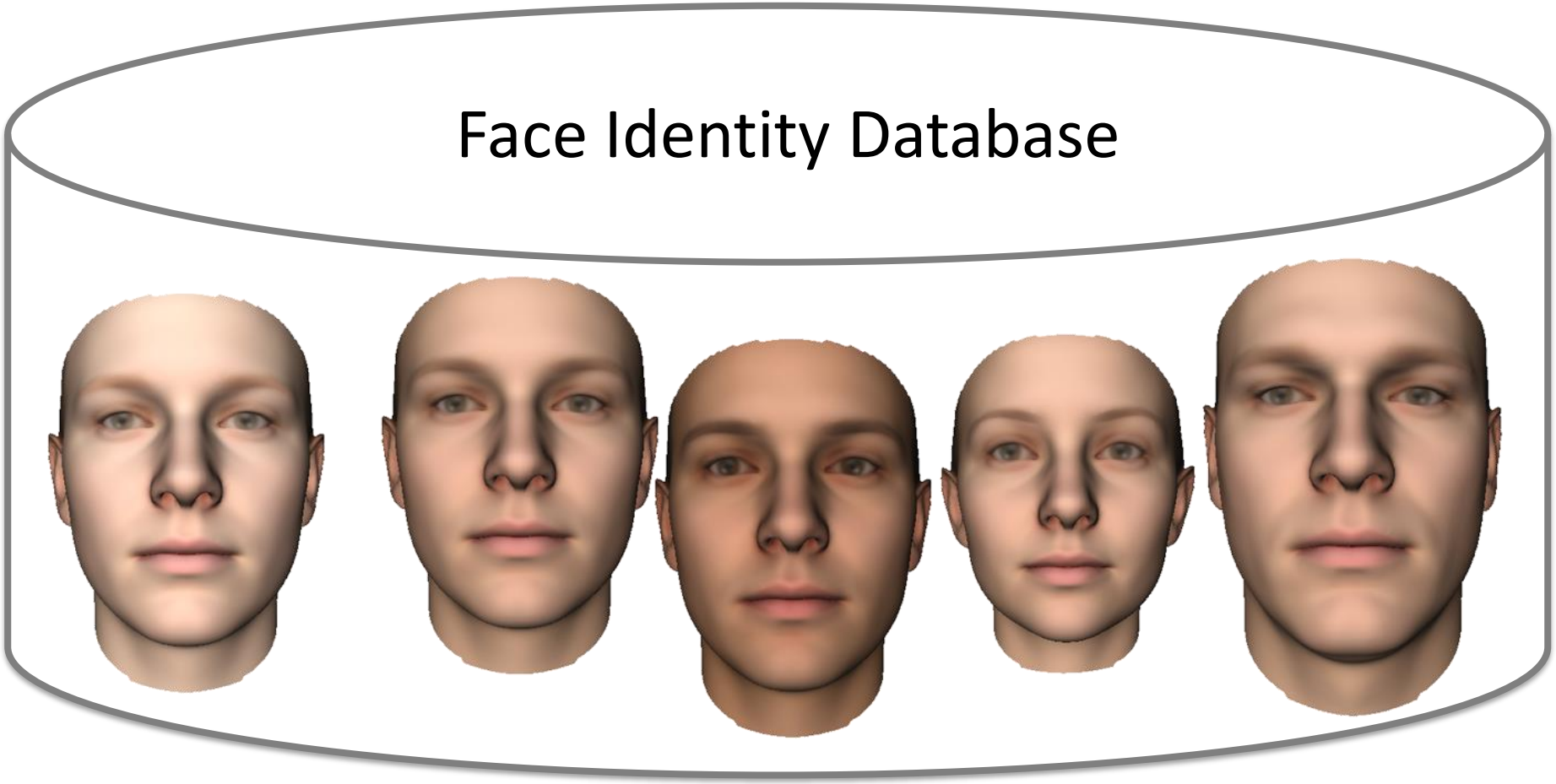
Reconstruction & Tracking of Non-rigid Scenes in *Real-Time*

Richard Newcombe, Dieter Fox, Steve Seitz

Computer Science and Engineering,
University of Washington

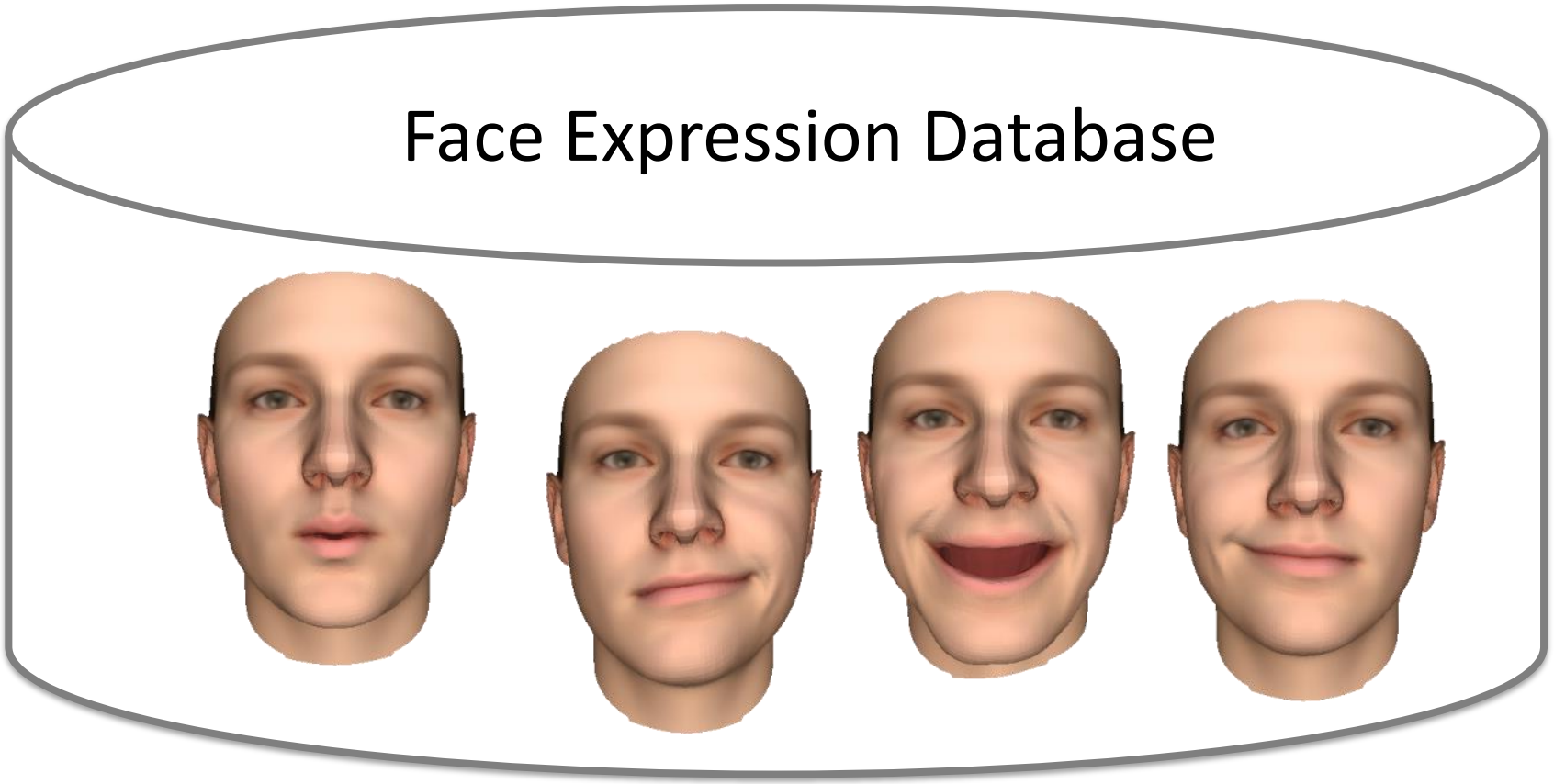
[Newcombe et al. 15]

Parametric Face Model: Shape Identity



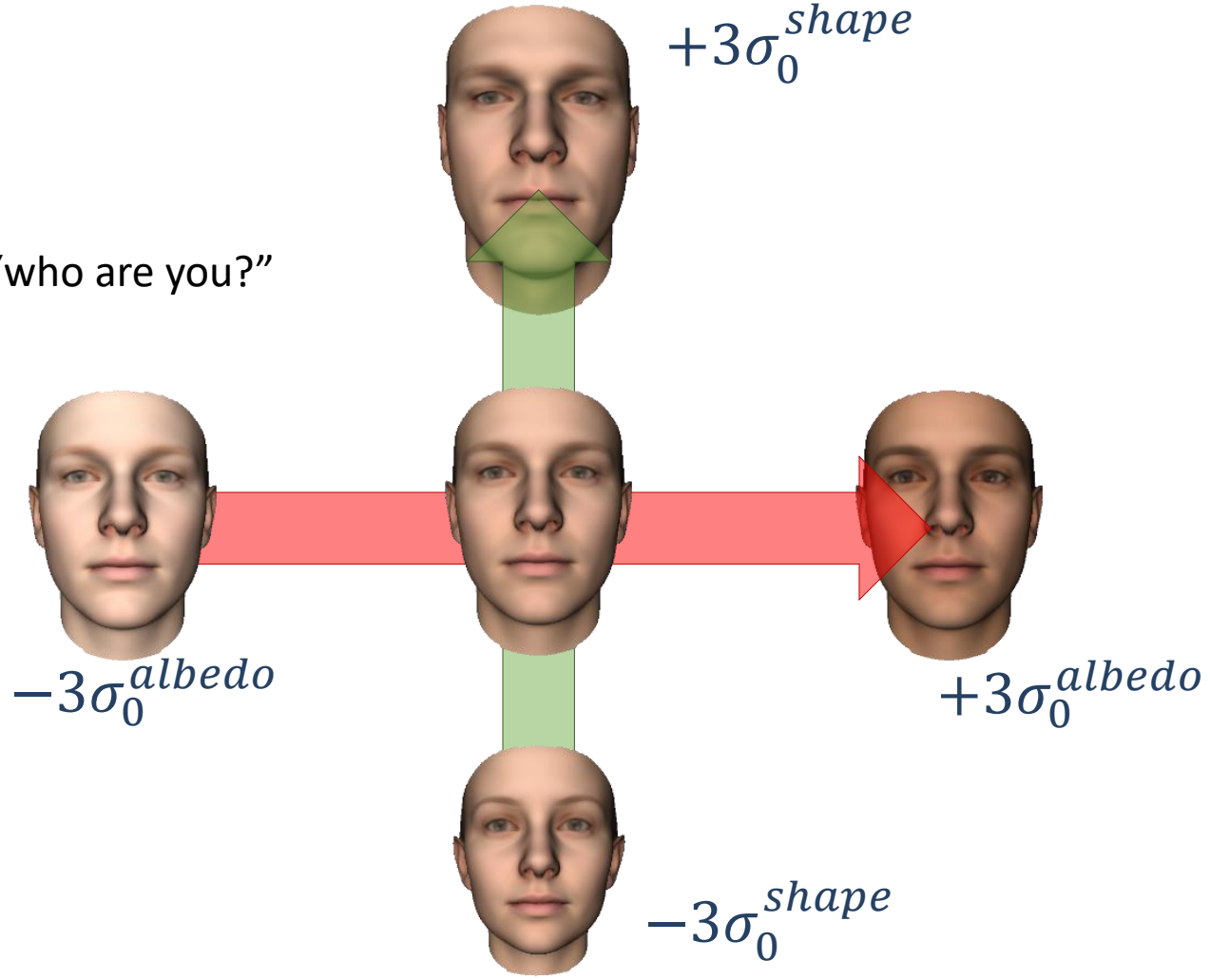
[Blanz and Vetter 99]: BlendShapes

Parametric Face Model



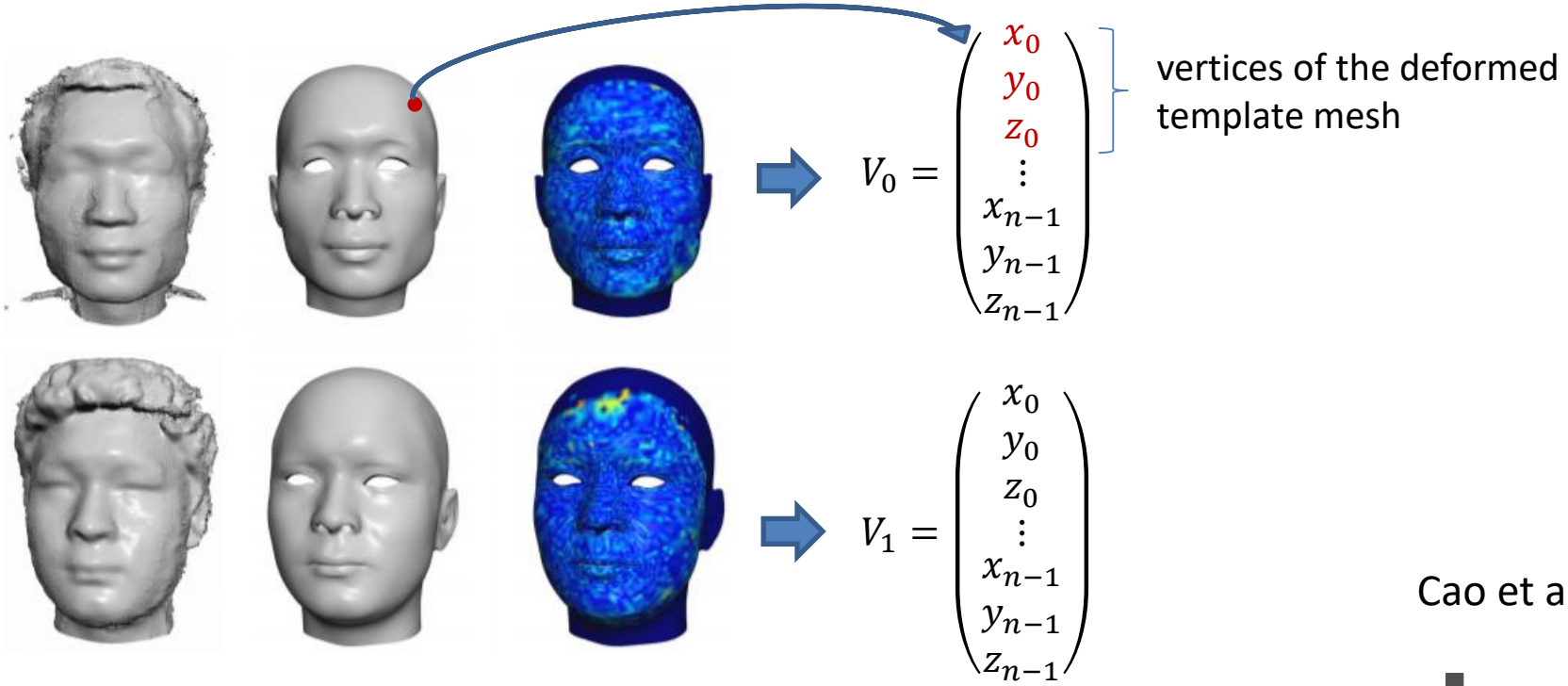
Parametric Face Model: Shape Identity

Shape Identity defines “who are you?”



How to Construct a Parametric Face Model?

- Compute face prior:
  - Scan a few hundred faces/expressions and fit a topologically-consistent template to each scan using non-rigid registration



Cao et al. “FaceWarehouse”

How to Construct a Parametric Face Model?

- Compute face prior:
  - Scan a few hundred faces/expressions and fit a topologically-consistent template to each scan using non-rigid registration
  - Compute PCA-basis for identity [and expressions]

$$V_0 = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \\ \vdots \\ x_{n-1} \\ y_{n-1} \\ z_{n-1} \end{pmatrix} \quad \cdots \quad V_{m-1} = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \\ \vdots \\ x_{n-1} \\ y_{n-1} \\ z_{n-1} \end{pmatrix}$$

PCA

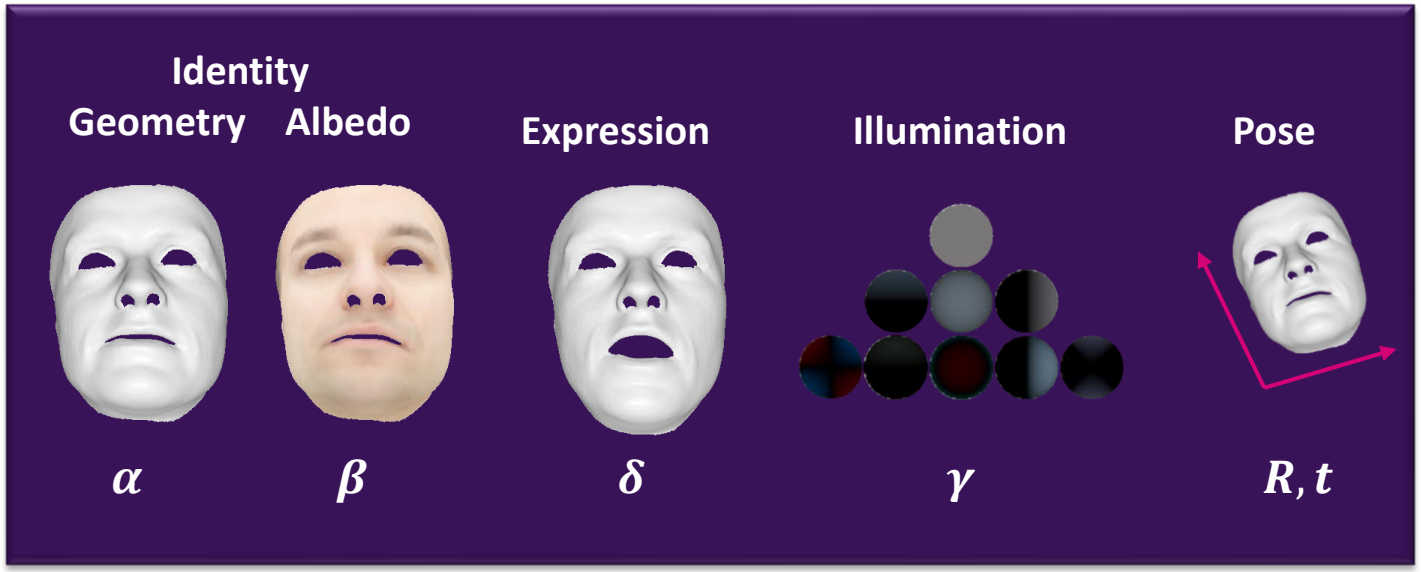
Average: $\bar{V}$

Principle Components: $\vec{E}_0 \ \cdots \vec{E}_{m-2}$

Std. Deviation: $\sigma_0 \ \cdots \sigma_{m-2}$

Face Model Fitting

$$P =$$



TOG'15 [Thies et al.]: Facical Reenactment

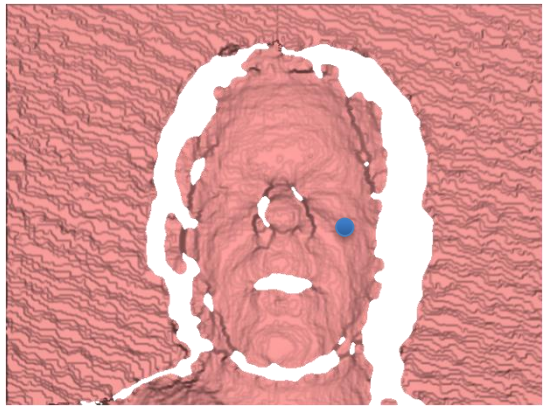
Face Tracking Objective

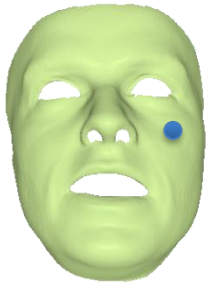
$$E(P) =$$

TOG'15 [Thies et al.]: Facical Reenactment

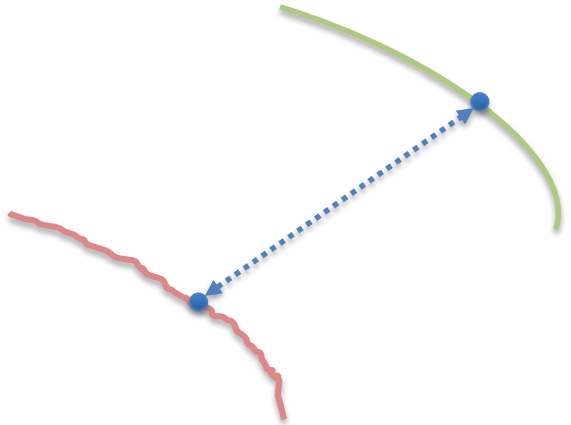
Face Tracking Objective

$$E(P) = \underbrace{E_{geom}(P)}_{\text{Geometry Consistency}}$$

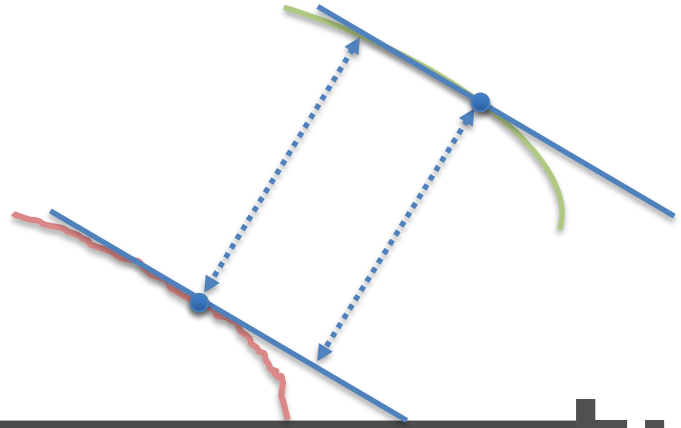




Point-to-Point



Symmetric Point-to-Plane



TOG'15 [Thies et al.]: Facical Reenactment

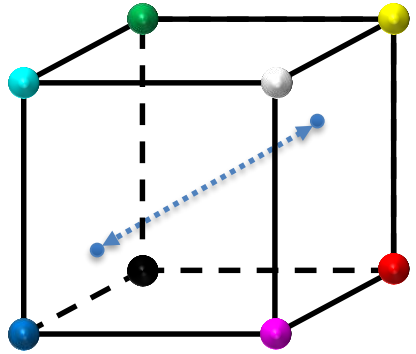
Face Tracking Objective

$$E(P) = E_{geom}(P) + E_{col}(P)$$

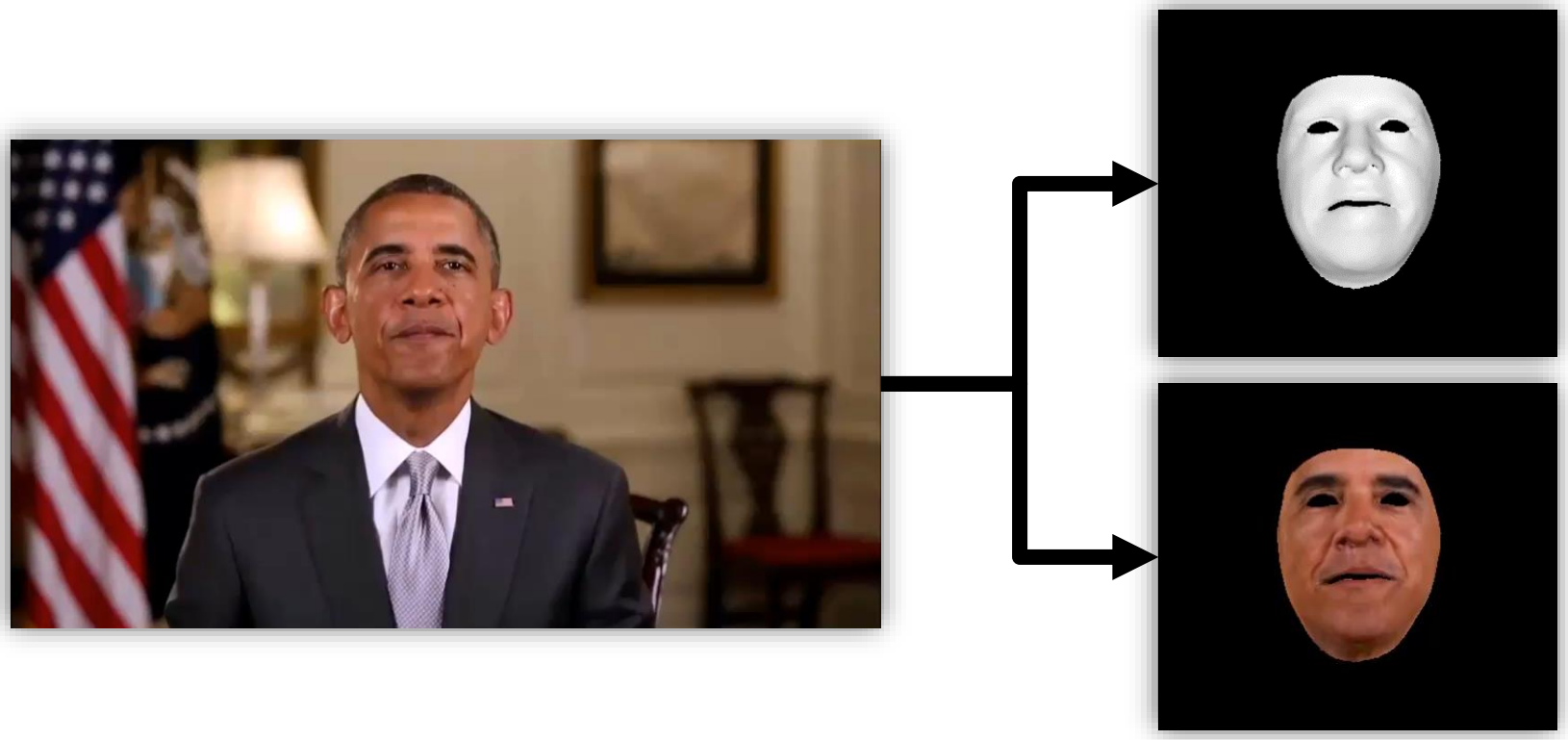
Geometry Consistency

Color Consistency

Distance in RGB Color Space



RGB-only Face Tracking



CVPR'16 [Thies et al.]: Face2Face

- Today: Body and Hand Tracking

Parametric Models

- Idea of parametric generalizes to other parts: hands and bodies

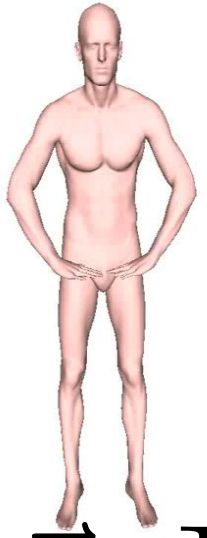
- Simple roadmap
  - Take set of scans from variety of subjects / poses
  - Non-rigidly align them with template to obtain consistent topology

- Often parameterization with respect to an underlying skeleton

Parametric Models: Virtual Humans

3D scan with texture

Ground truth shape



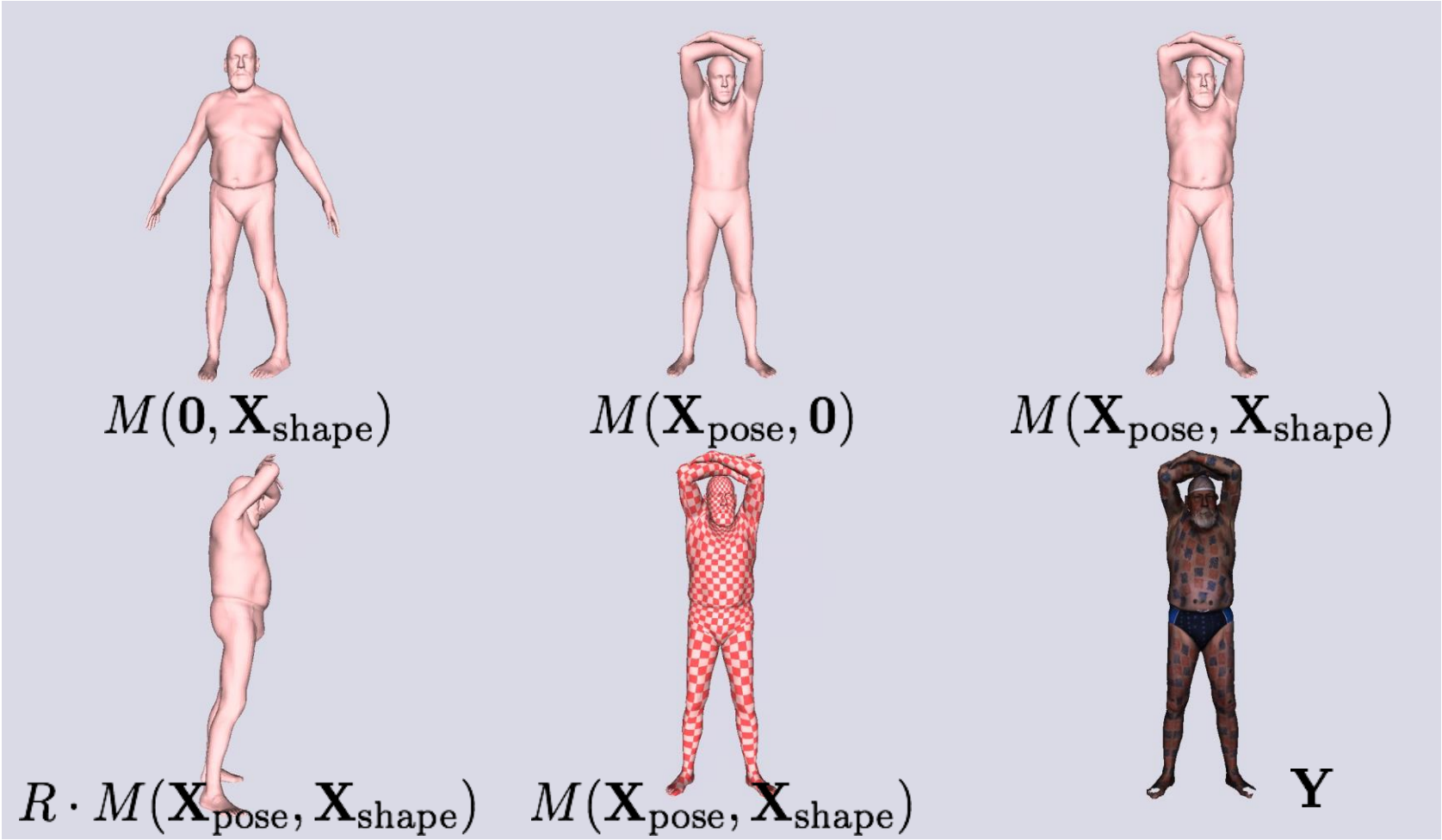
Model

Model with texture

$$M(\vec{\theta}, \vec{\beta}, \vec{u}; \Phi) \longrightarrow \text{3D mesh}$$

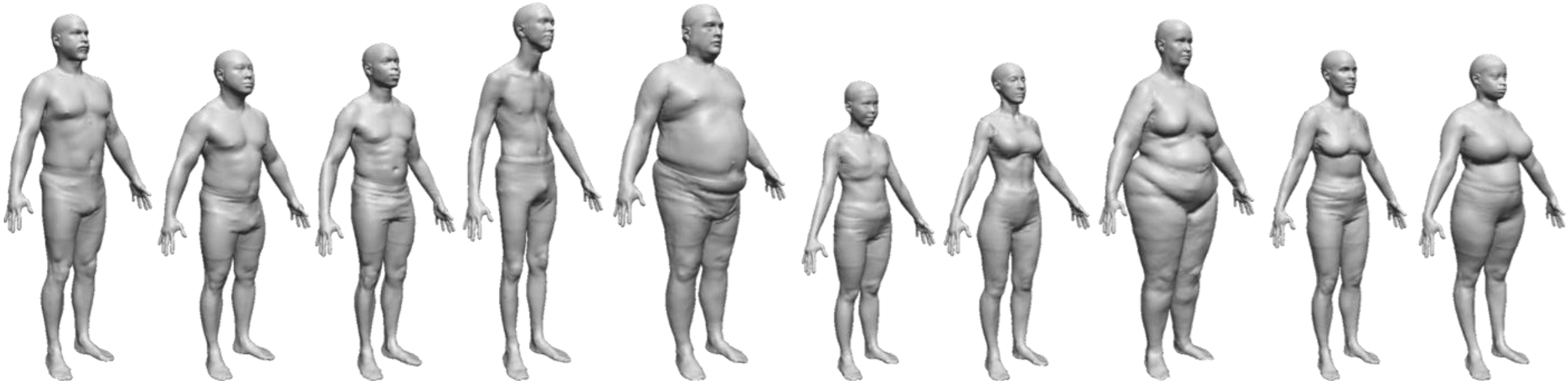
pose shape texture Hyper-parameters to learn

A Body Model is a Function



Modeling the Space of Human Bodies

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

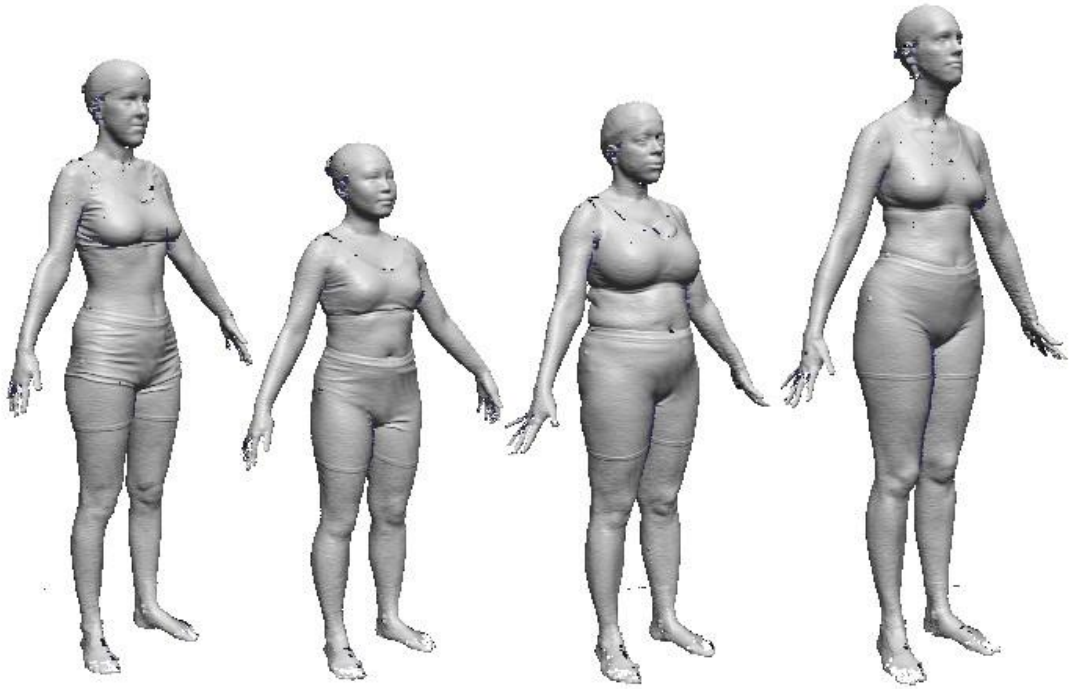


[Allen et al. 03]

Modeling the Space of Human Bodies

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

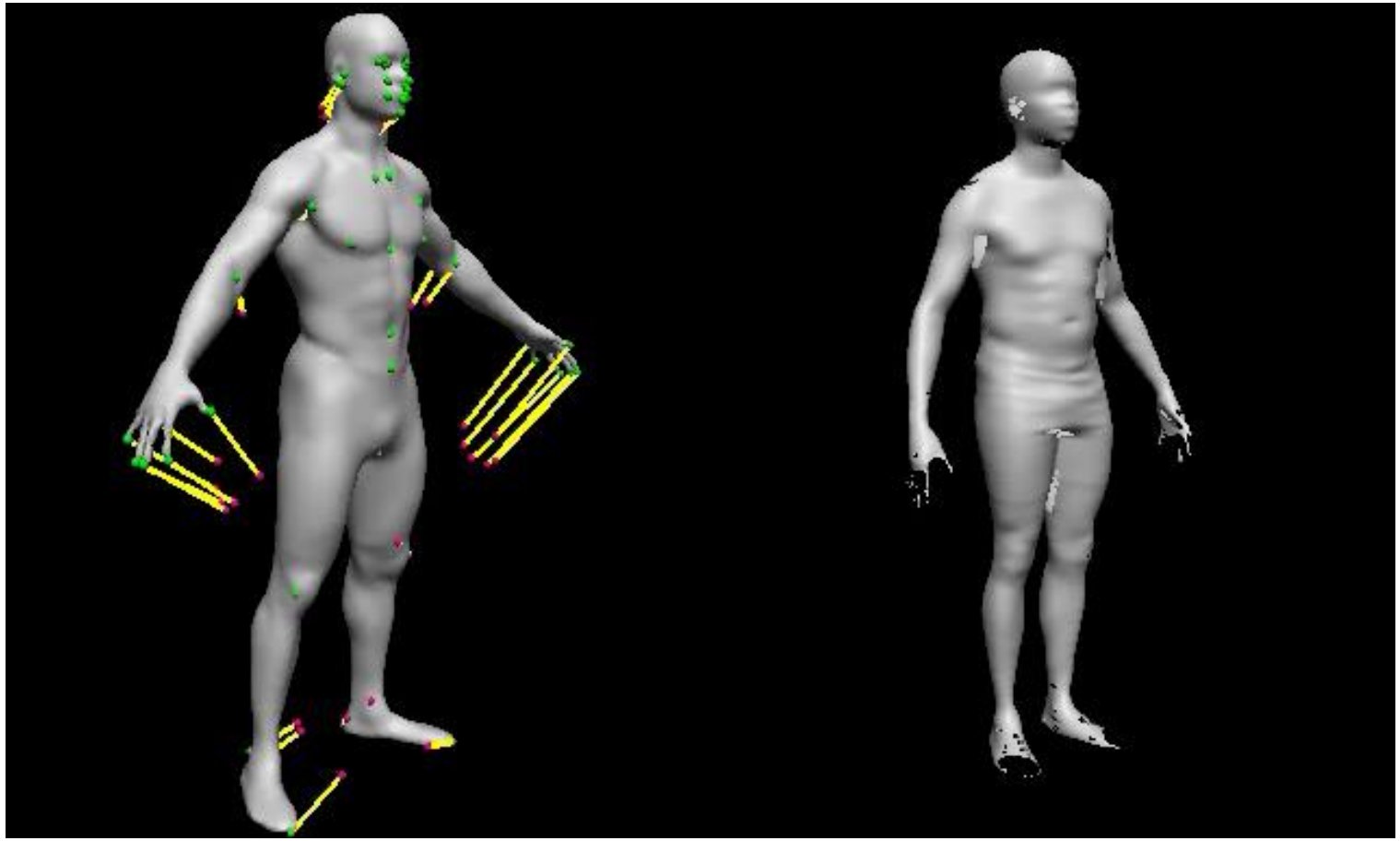
Cyberware 3D whole body scanner



[Allen et al. 03] Slide V. Kim

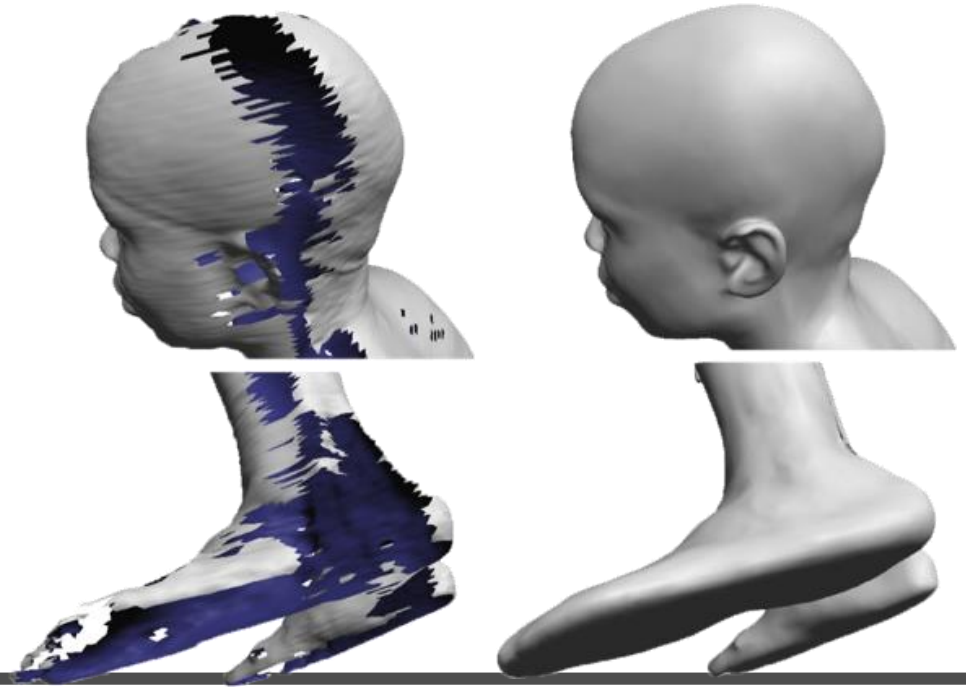
Modeling the Space of Human Bodies

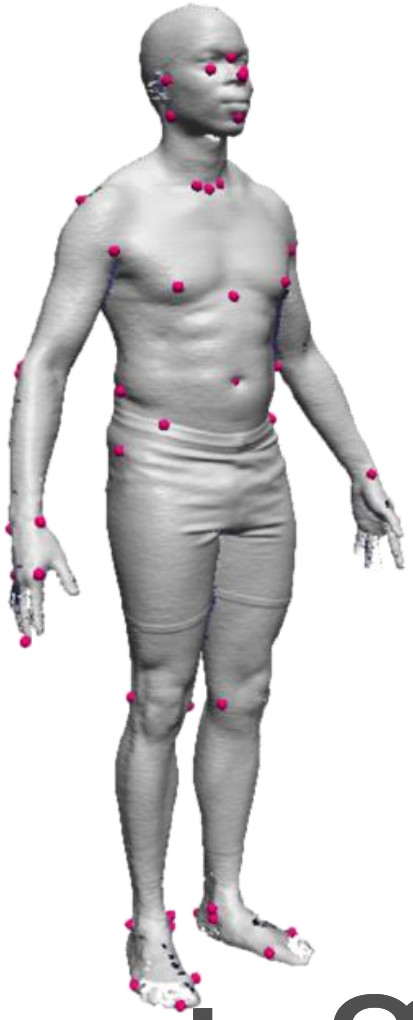
- Non-rigid registration to align scans



Modeling the Space of Human Bodies

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

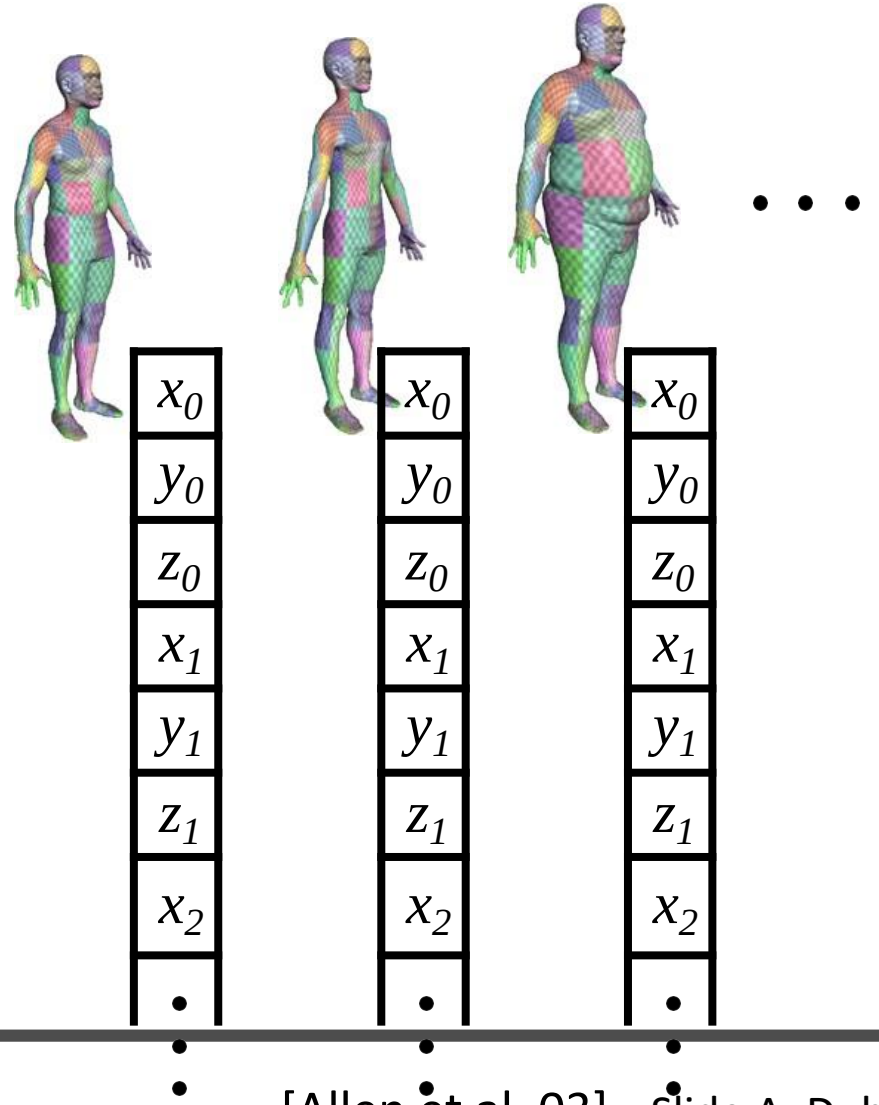




[Allen et al. 03] Slide A. Dubrovina

Variance in Human Bodies

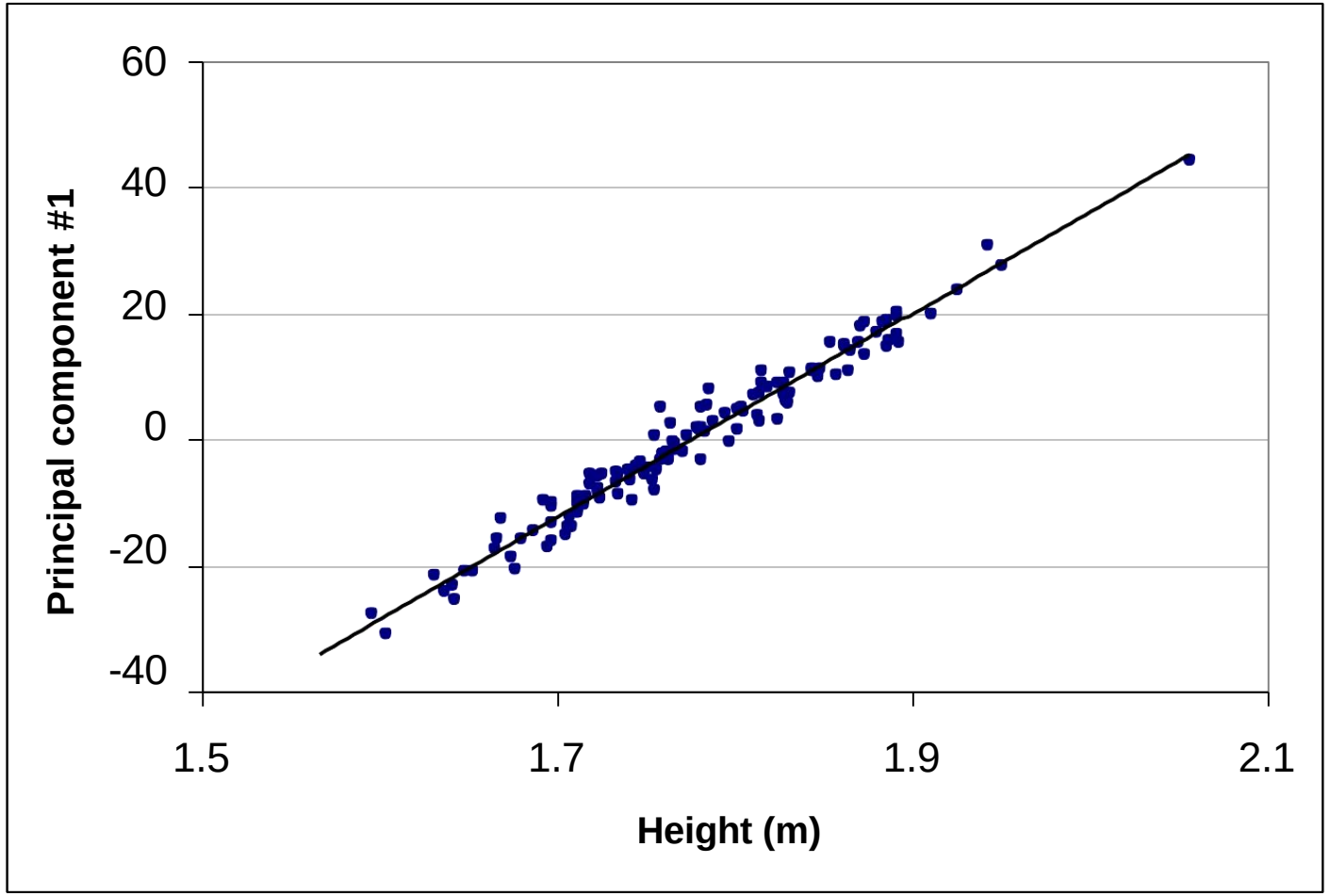
Compute statistical model using PCA



average male

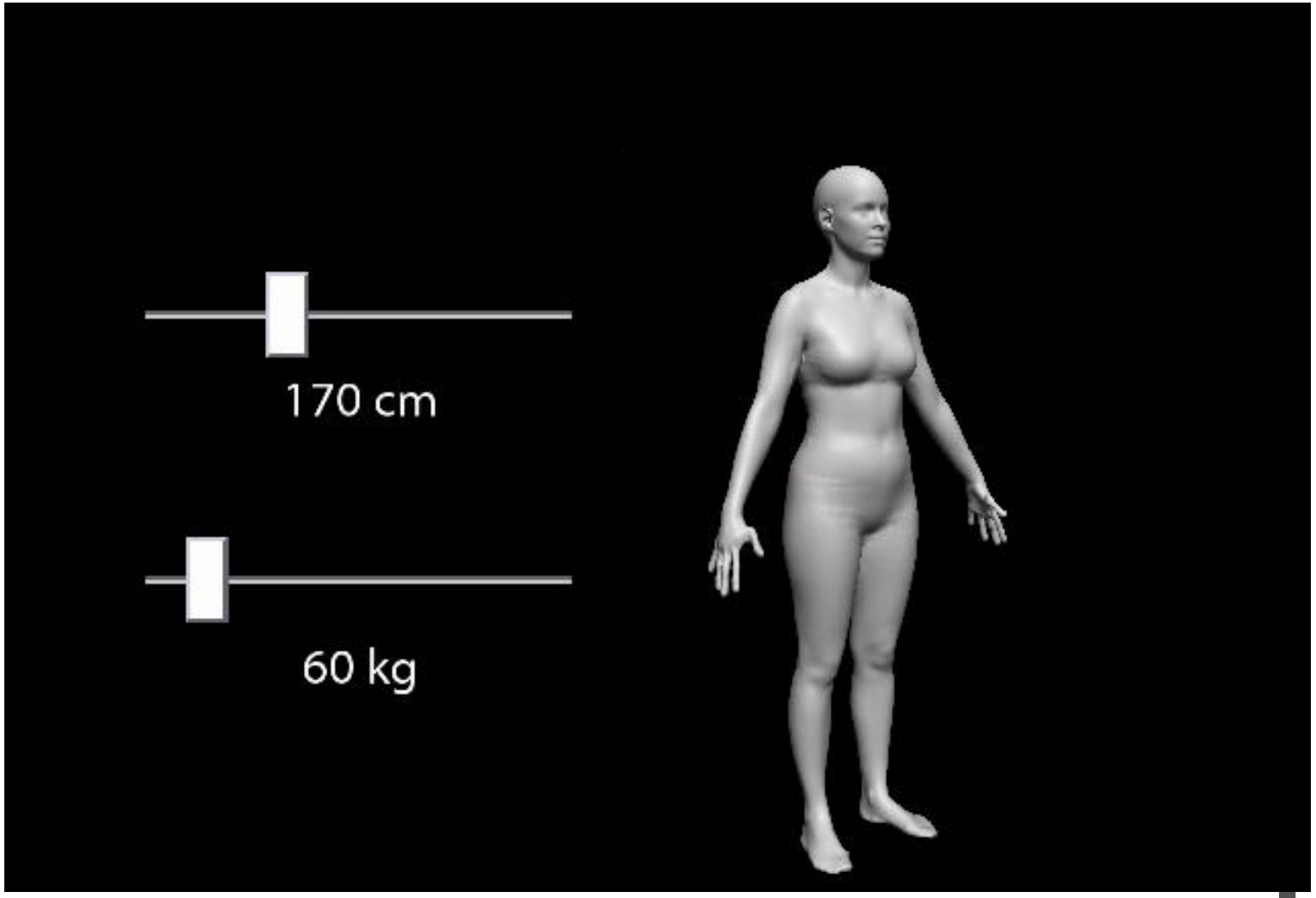
[Allen et al. 03] Slide A. Dubrovina

Variance in Human Bodies



[Allen et al. 03] Slide A. Dubrovina

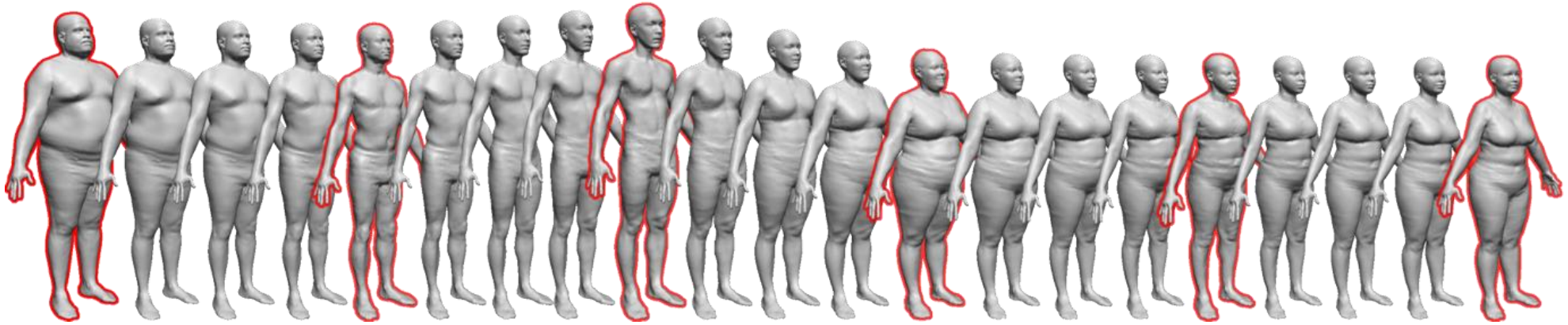
Variance in Human Bodies



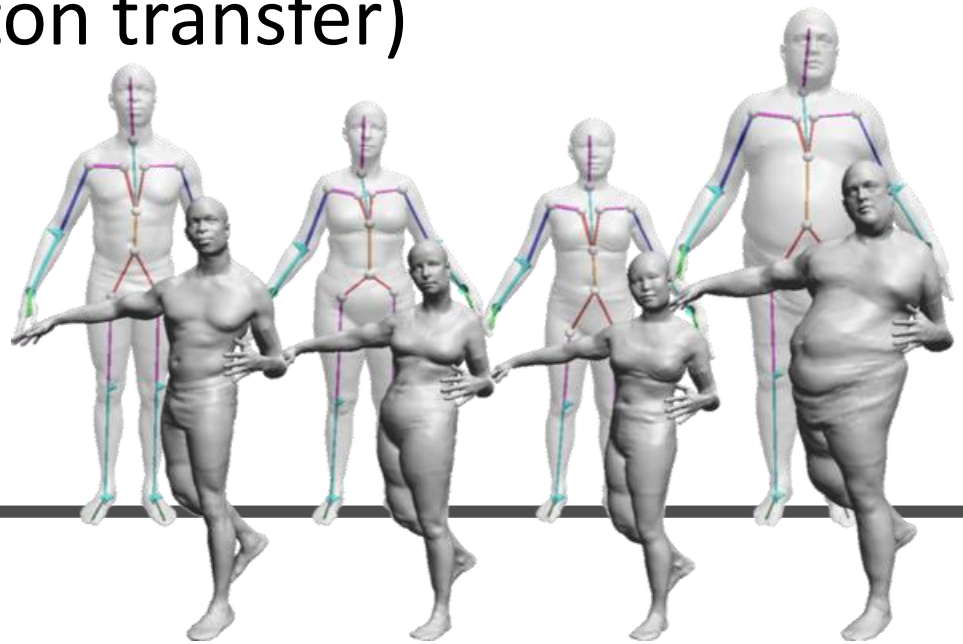
[Allen et al. 03]

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

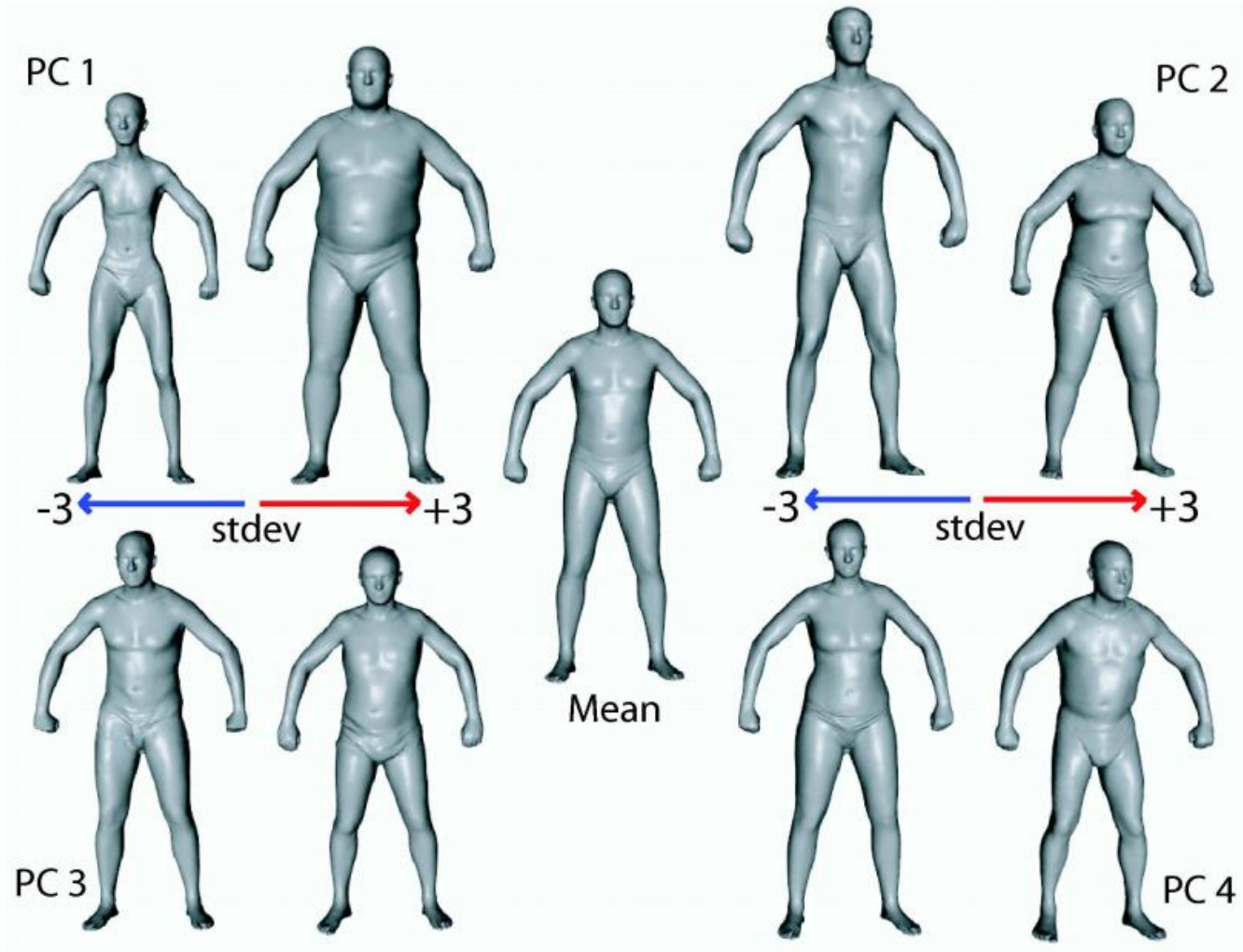
- Poses: 70 poses of a particular person in a wide variety of poses
- Identity: 37 different people in a similar (bot not identical) pose

+ CEASAR

- Reconstruct skeleton (16 parts) for each pose
- Computer Delta-‘PCA’-like blendshapes over triangles wrt. skel.

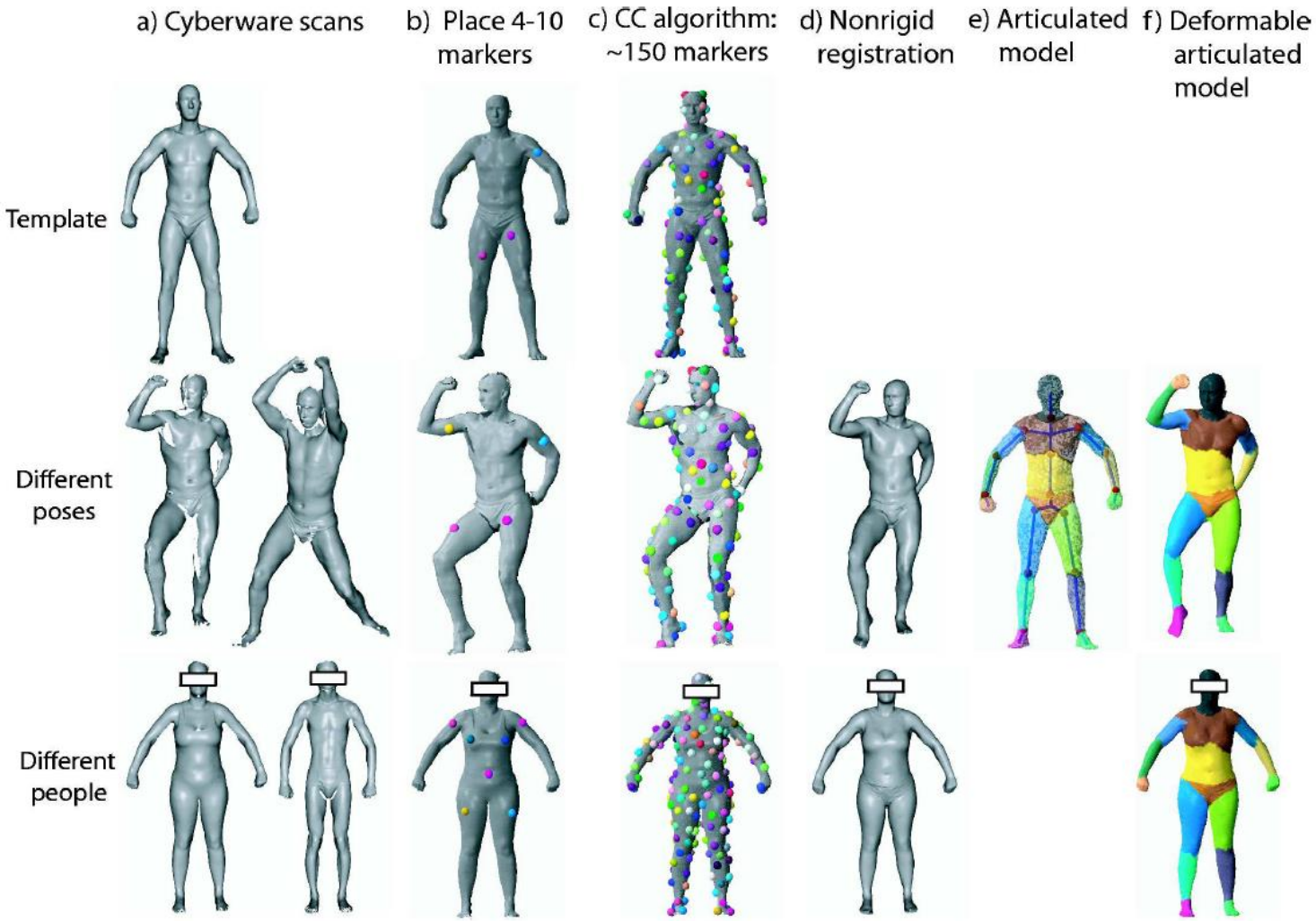
[Anguelov et al. 05]

Principle Components of SCAPE



[Anguelov et al. 05]

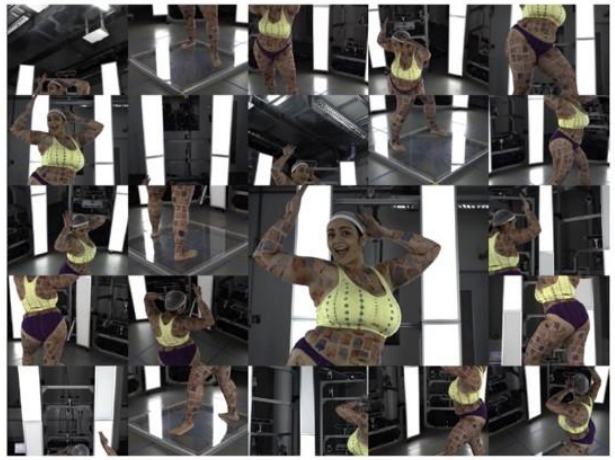
SCAPE: Shape Completion and Animation of People



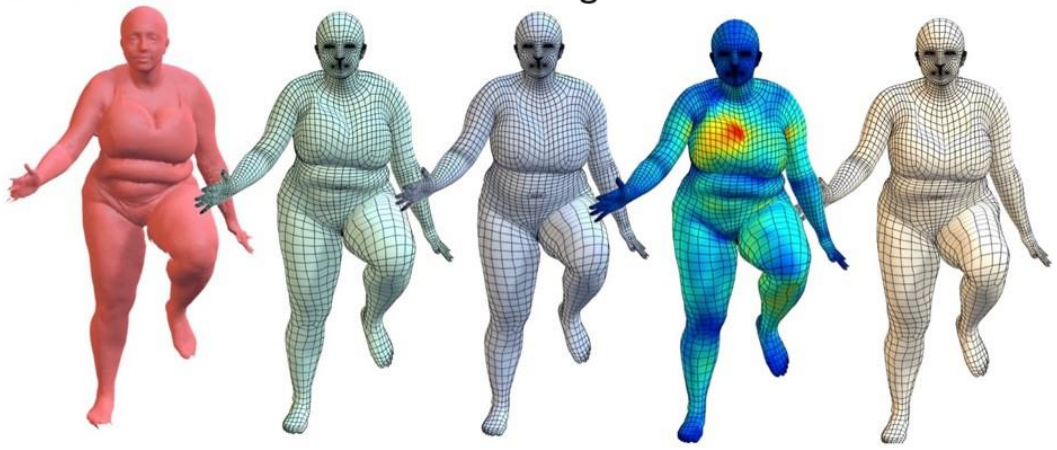
[Anguelov et al. 05]

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

Breathing model



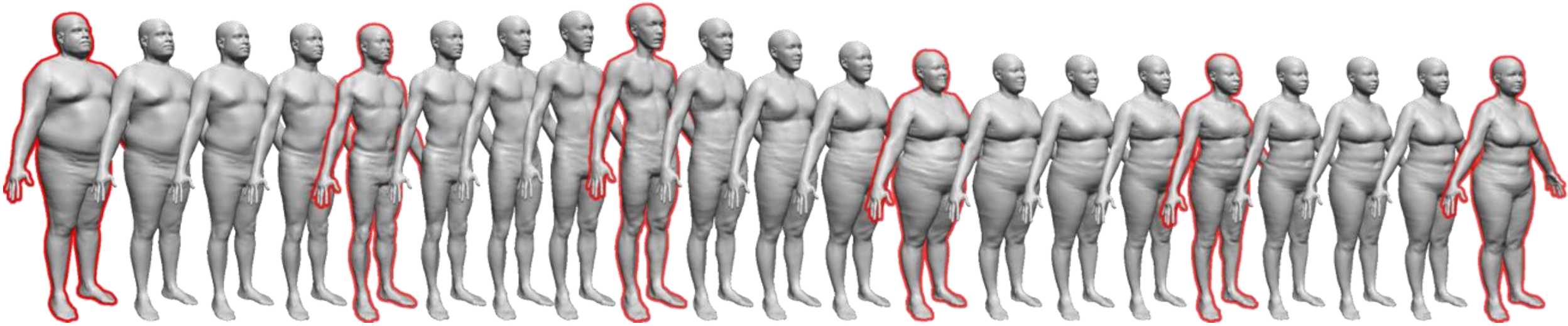
4D scanner (60 fps)



4D scan 4Cap BlendSCAPE Soft tissue Dyna

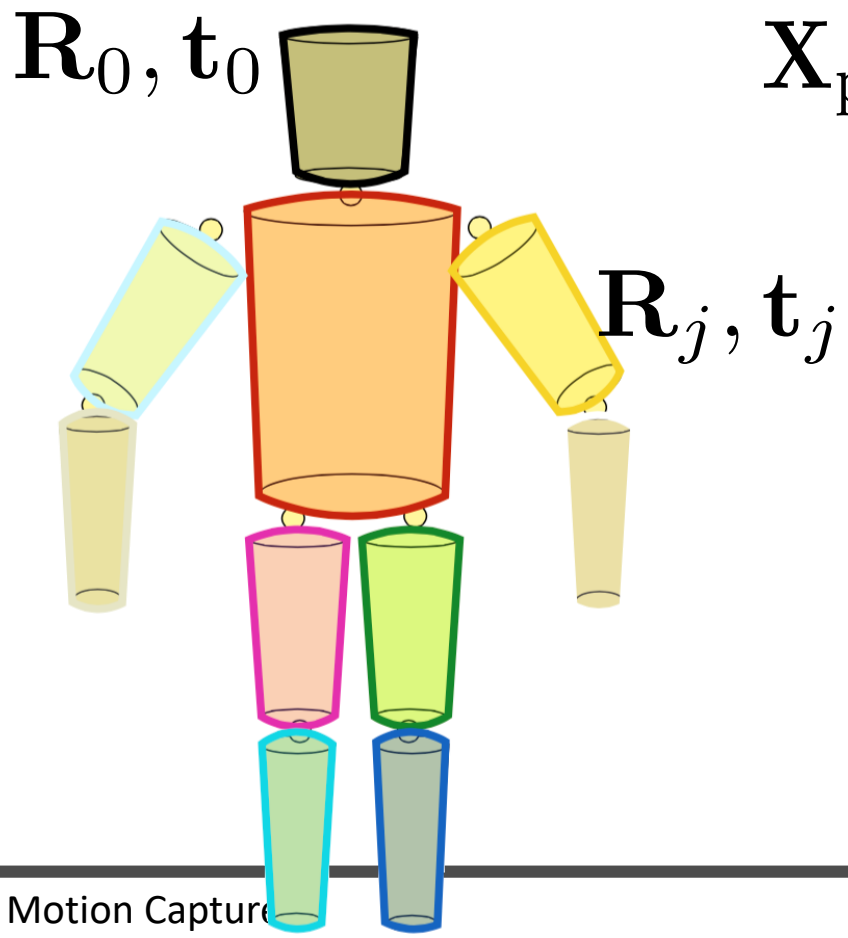
Parametrization of Shape

- PCA / Blendshapes



How do we parameterize pose ?

- Parameterize every body part separately ?



$$\mathbf{X}_{\text{pose}} = \{\mathbf{R}_0, \mathbf{t}_0, \ldots \mathbf{R}_N, \mathbf{t}_N\}$$

Problems ?

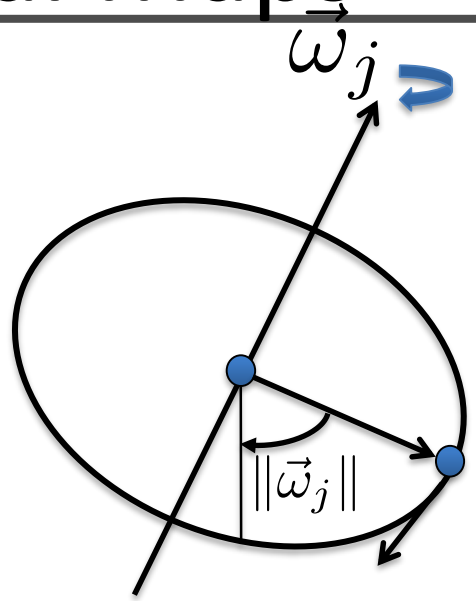
Rotation Parameterization

- 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

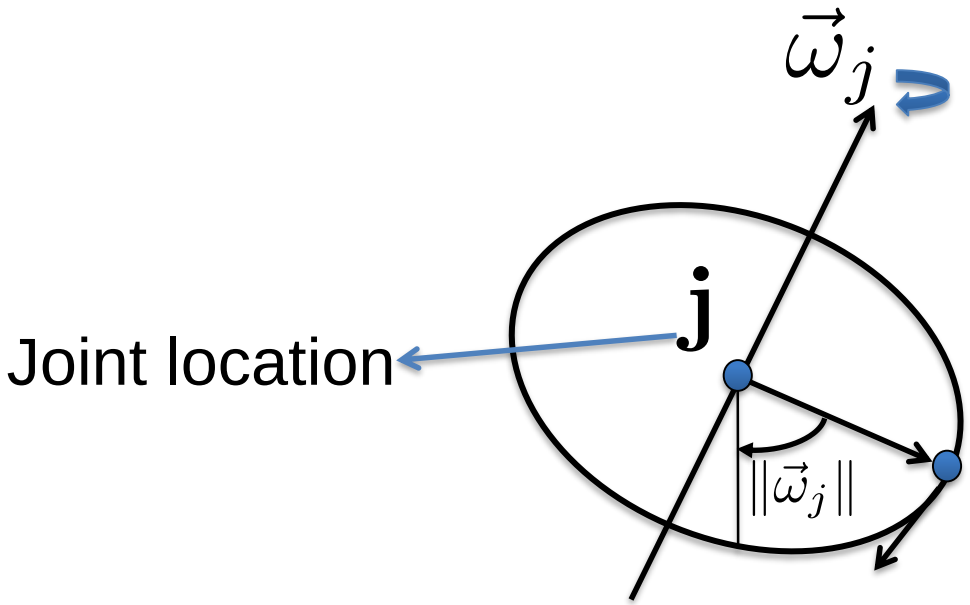


Rotation obtained with Rodrigues formula:

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

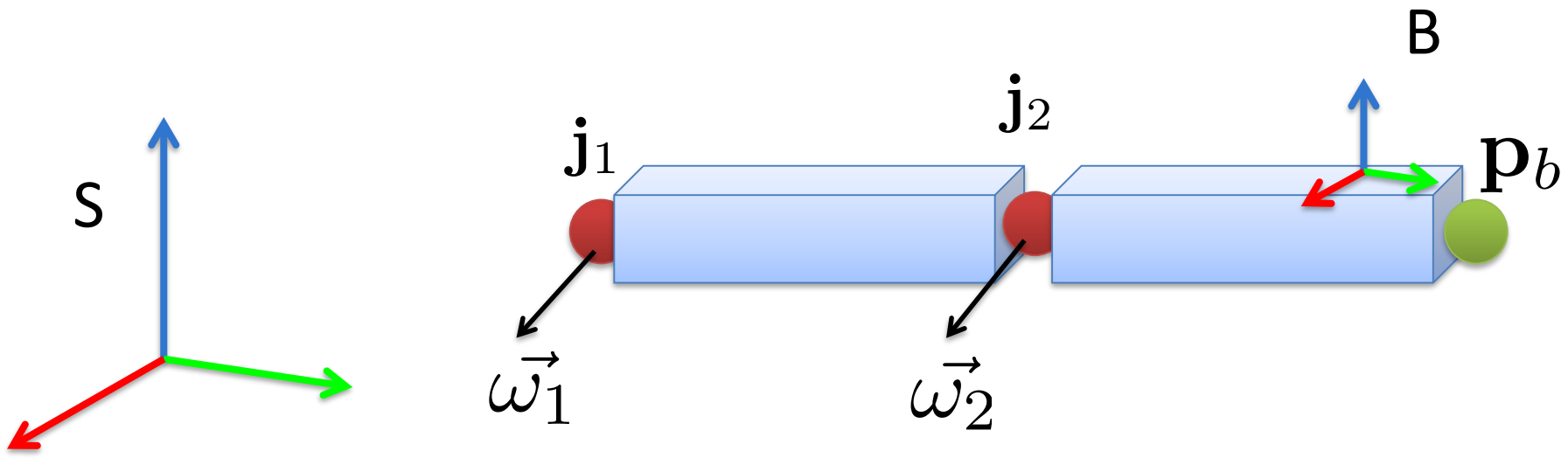
Joint Rigid Body Motion

The transformation associated with a rotational joint is

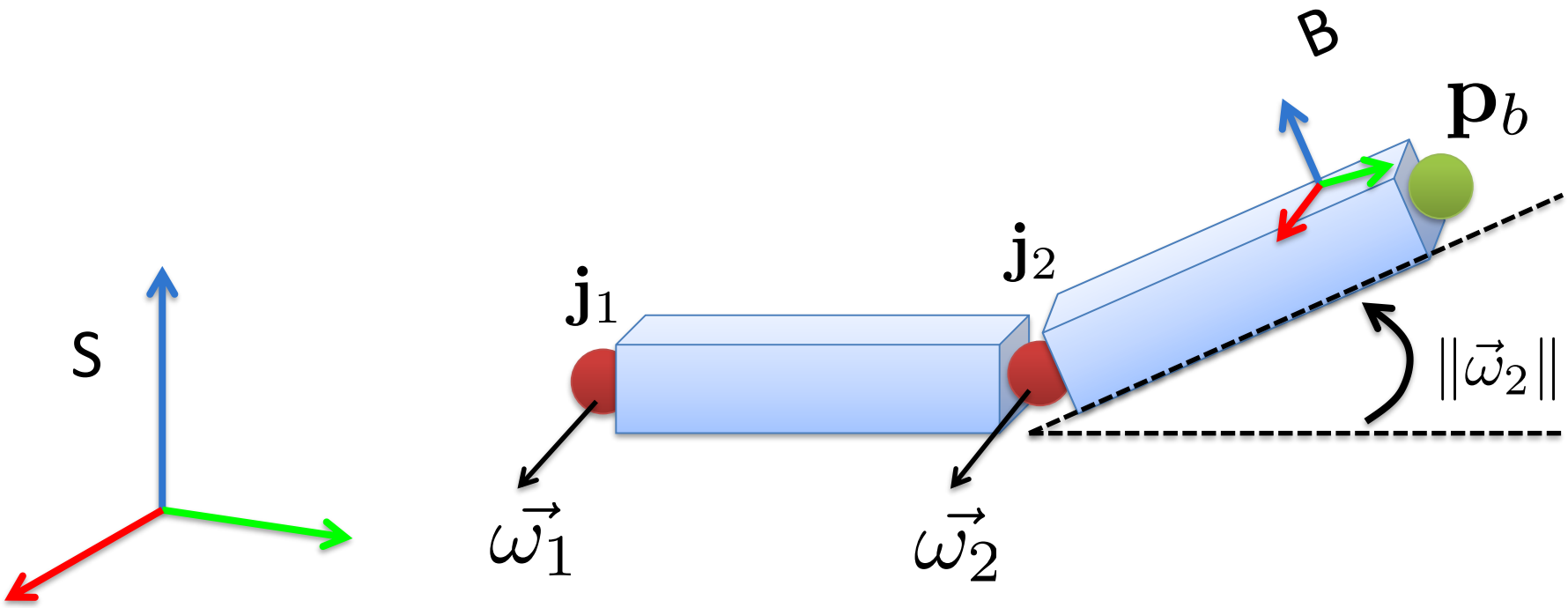


$$G(\vec{\omega}, \mathbf{j}) = \begin{bmatrix} [e^{\vec{\omega}}]_{3\times3} & \mathbf{j}_{3\times1} \\ \mathbf{0}_{1\times3} & 1 \end{bmatrix} \longrightarrow$$ Rigid Body Motion

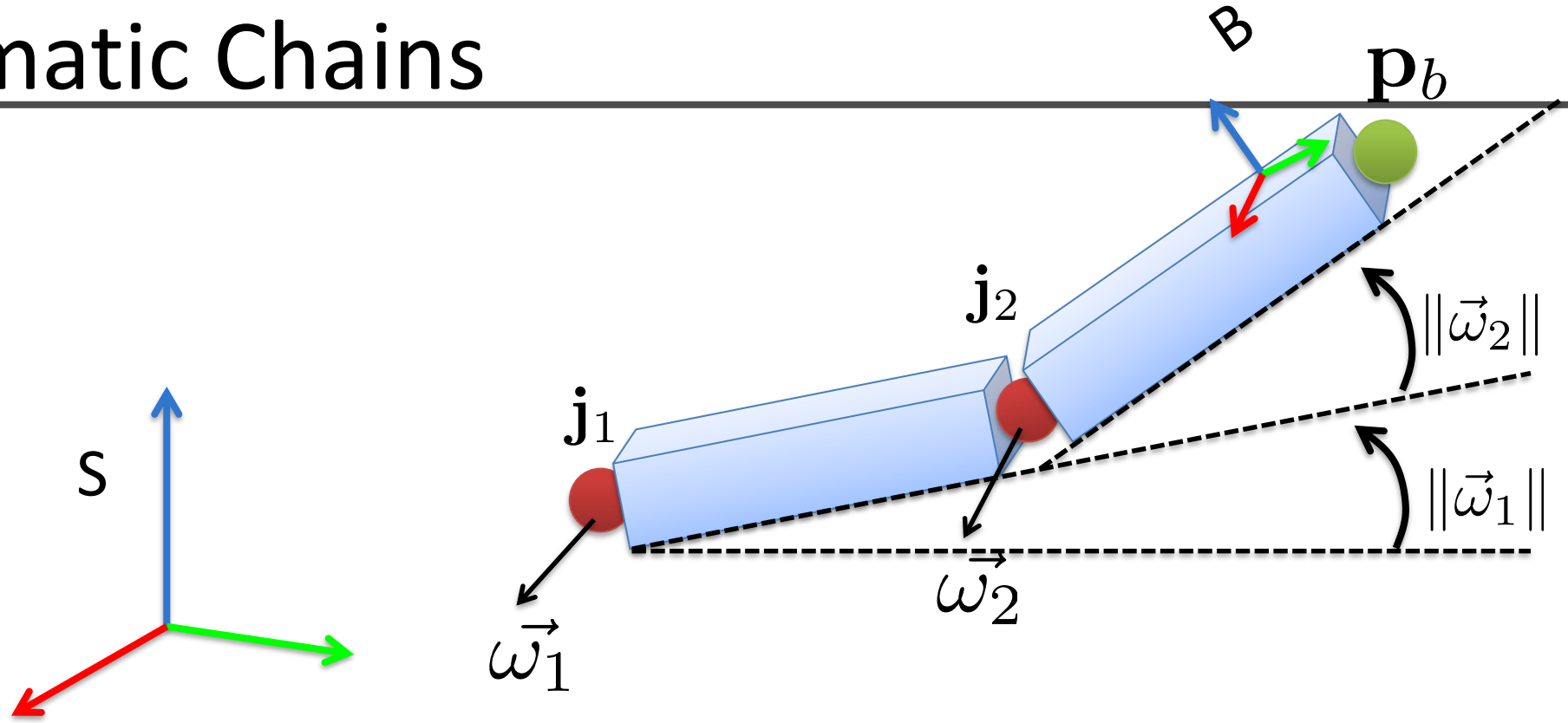
Kinematic Chains



Kinematic Chains



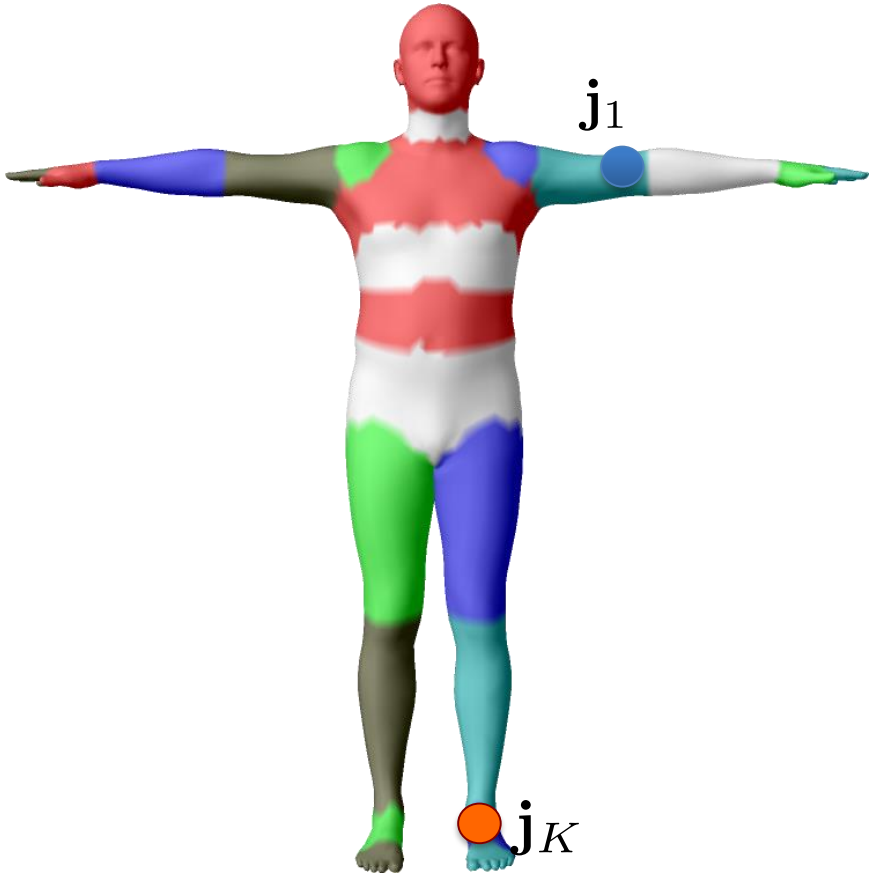
Kinematic Chains



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



$\mathbf{T}$

Given a set of joint locations

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

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

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

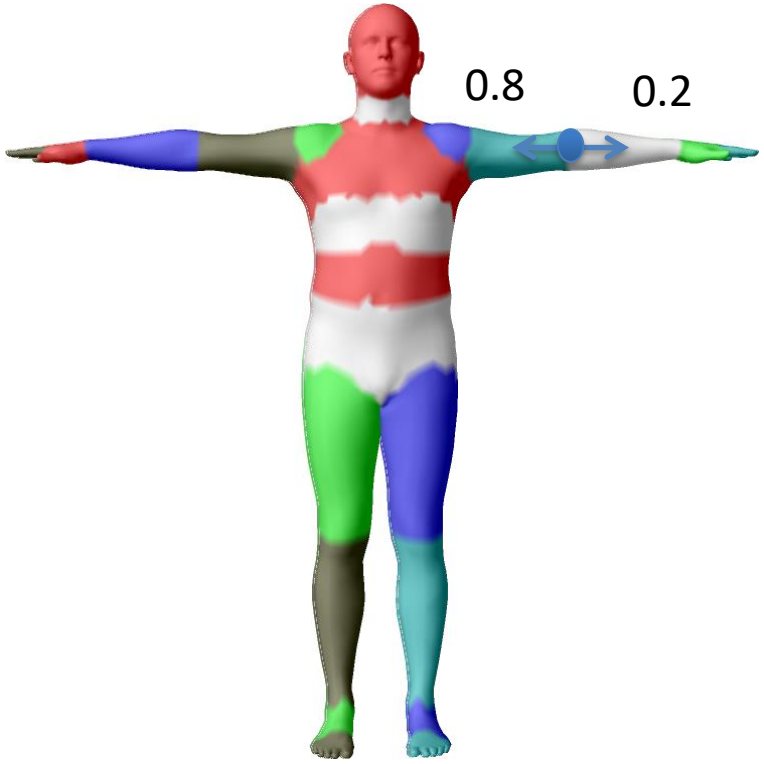
Pons-Moll & Rosenhahn 2011
Model-based Pose Estimation. Looking at People.

Linear Blend Skinning

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

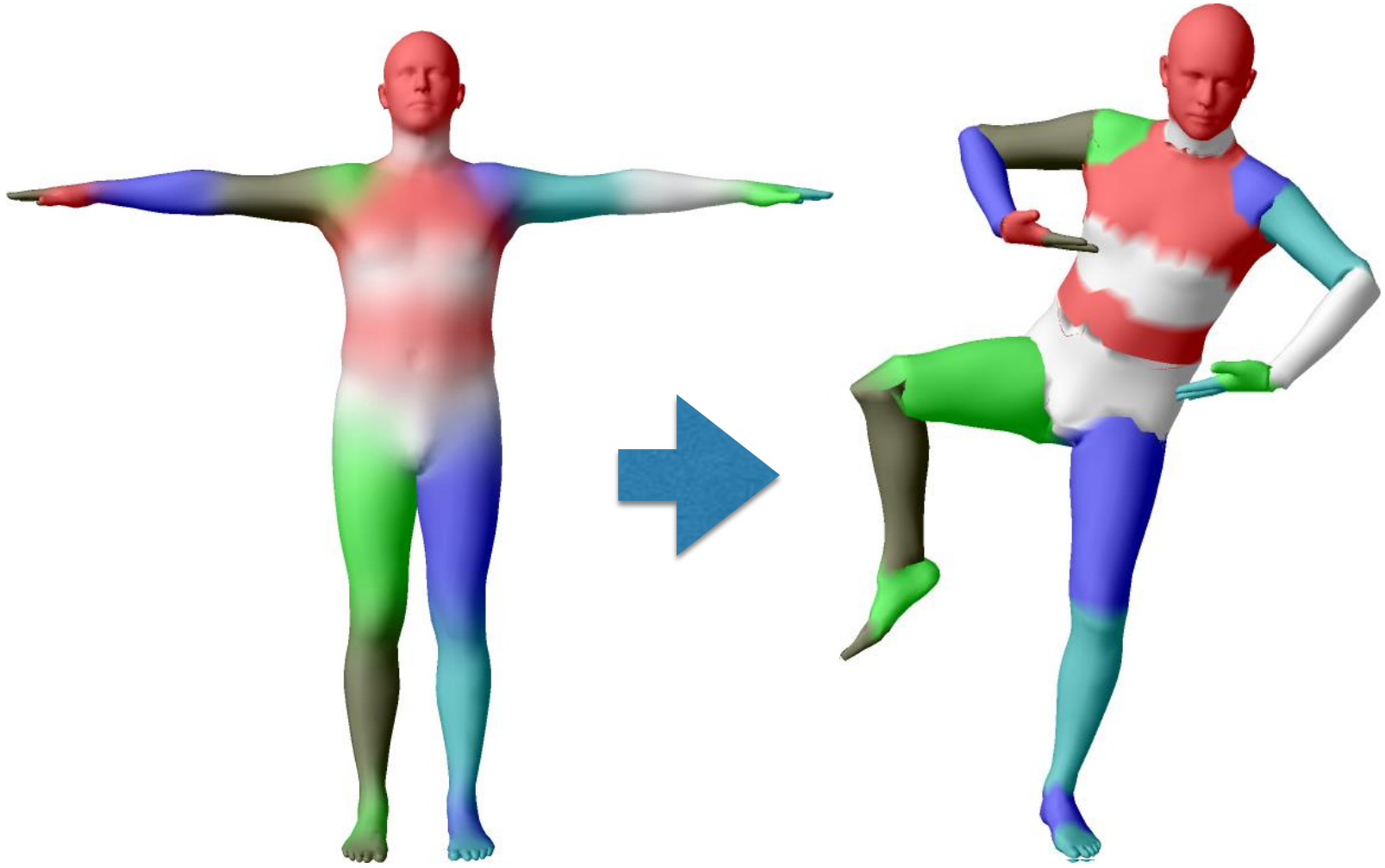
Blend weights

Part transformations



Points transformed as blended linear combination of joint transformation matrices

Linear Blend Skinning



Standard Skinning

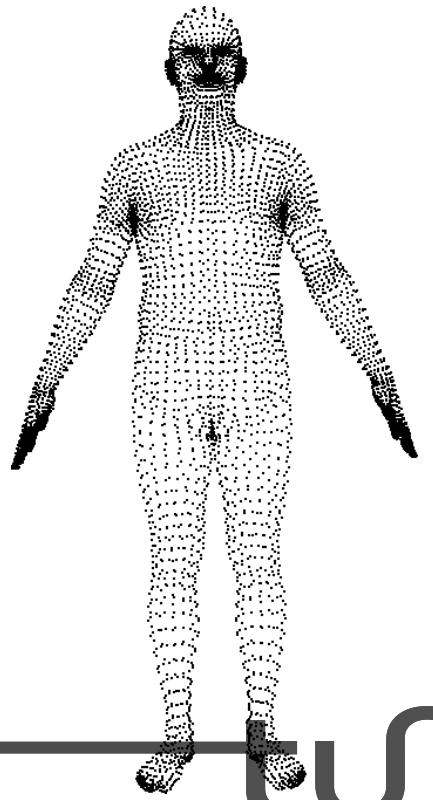
Standard skinning produces vertices from…

– 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

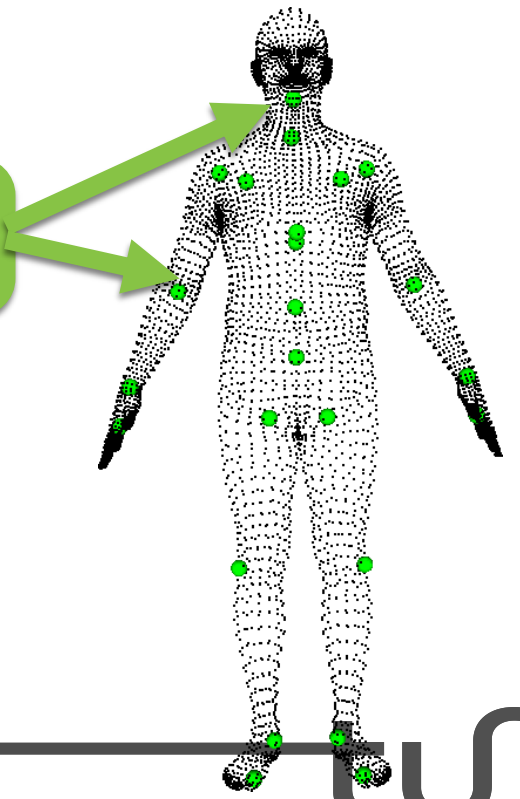
Standard skinning produces vertices from…

– 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

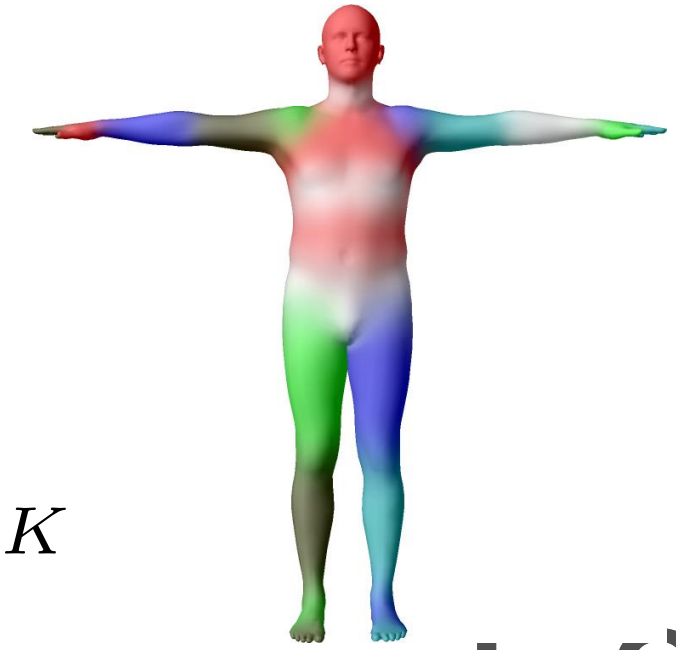
Standard skinning produces vertices from…

– 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

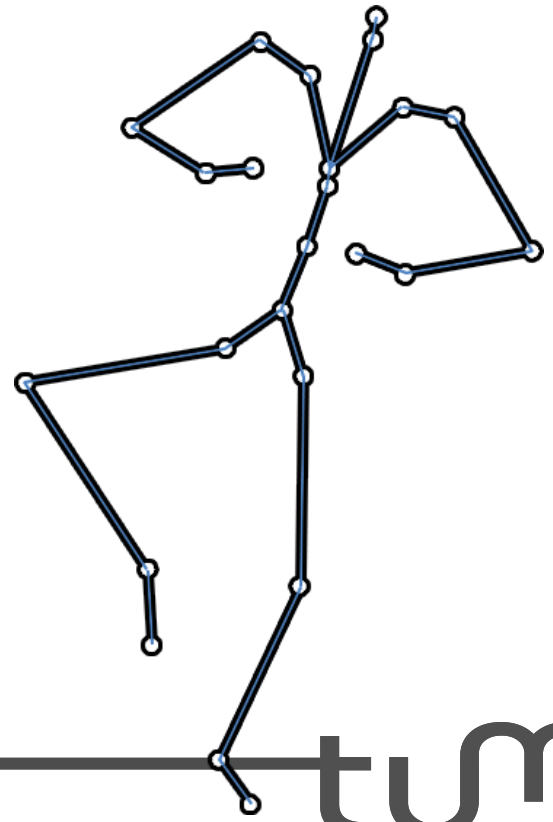
Standard skinning produces vertices from…

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

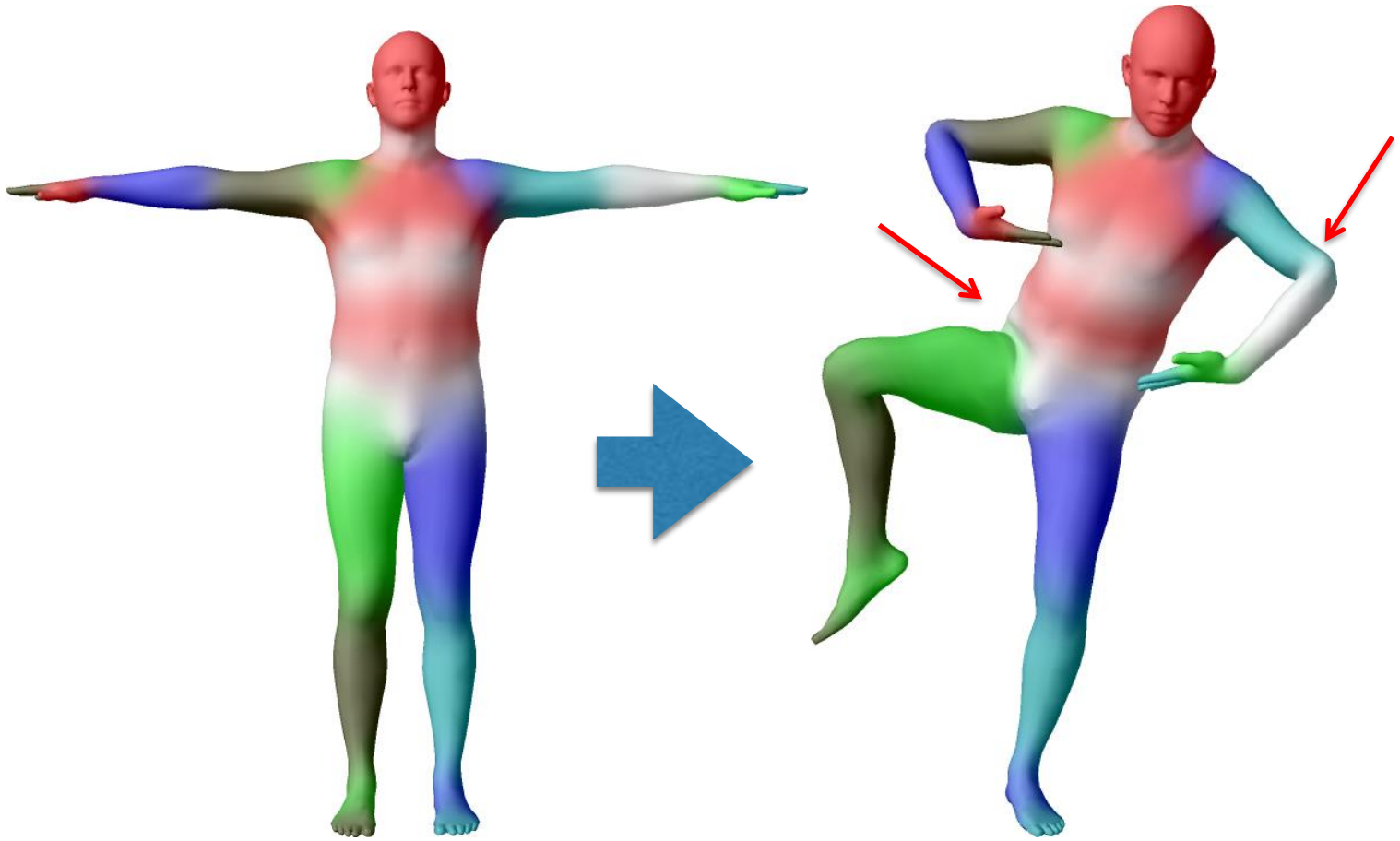


Skinning Function

- 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

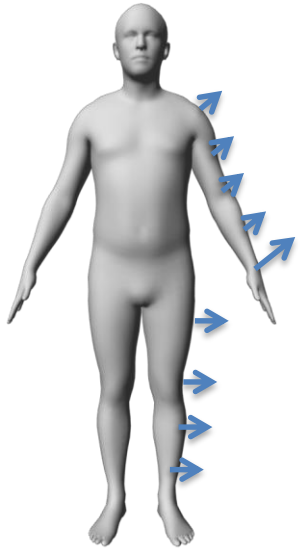


Slide G. Pons-Moll

Solution: (Corrective) Blend Shapes

- A **blend shape** is a set of vertex displacements in a rest pose

  – Pose blend shapes: correct for LBS problems

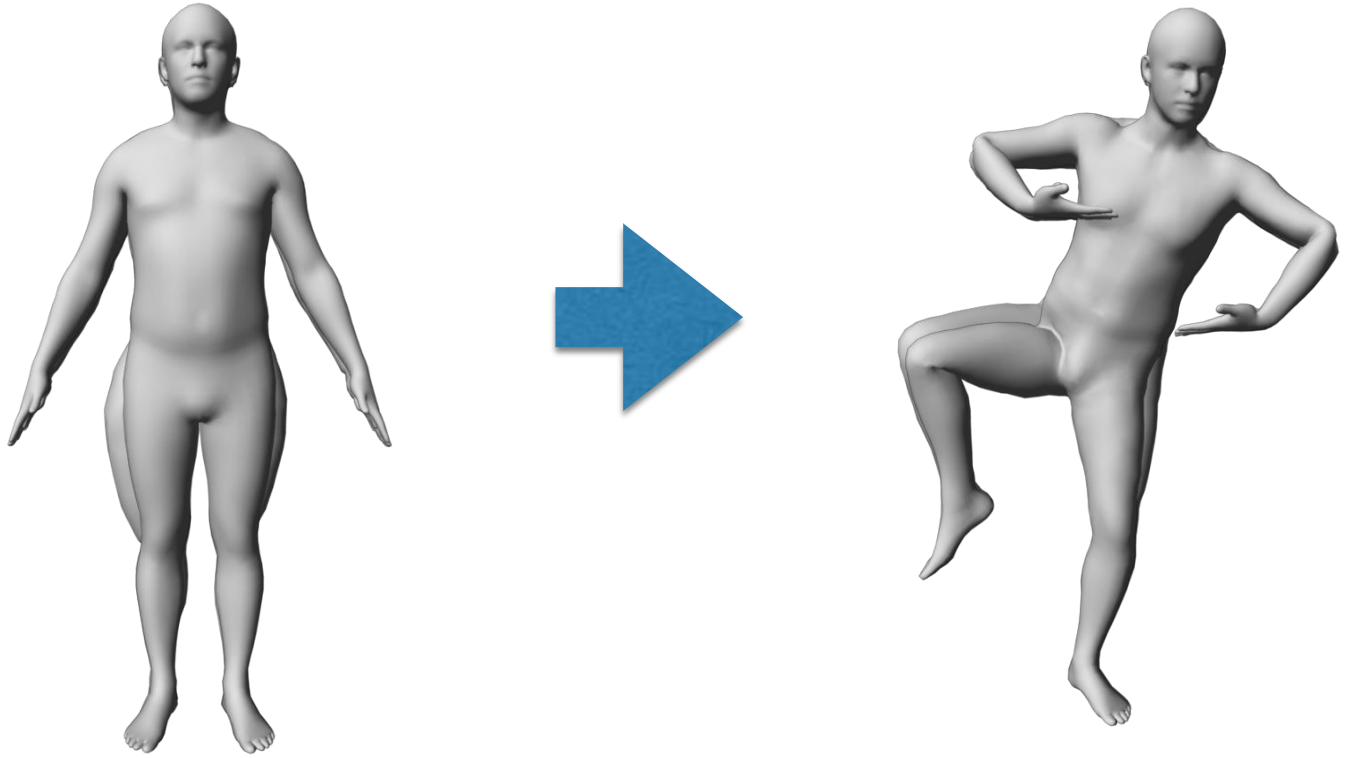


$$\mathbf{P} = \mathrm{vec}\left(\begin{bmatrix} \Delta x_1 & \Delta y_1 & \Delta z_1 \\ & \vdots & \\ \Delta x_N & \Delta y_N & \Delta z_N \end{bmatrix} \right) \in \mathbb{R}^{3N}$$

→ Offset 1

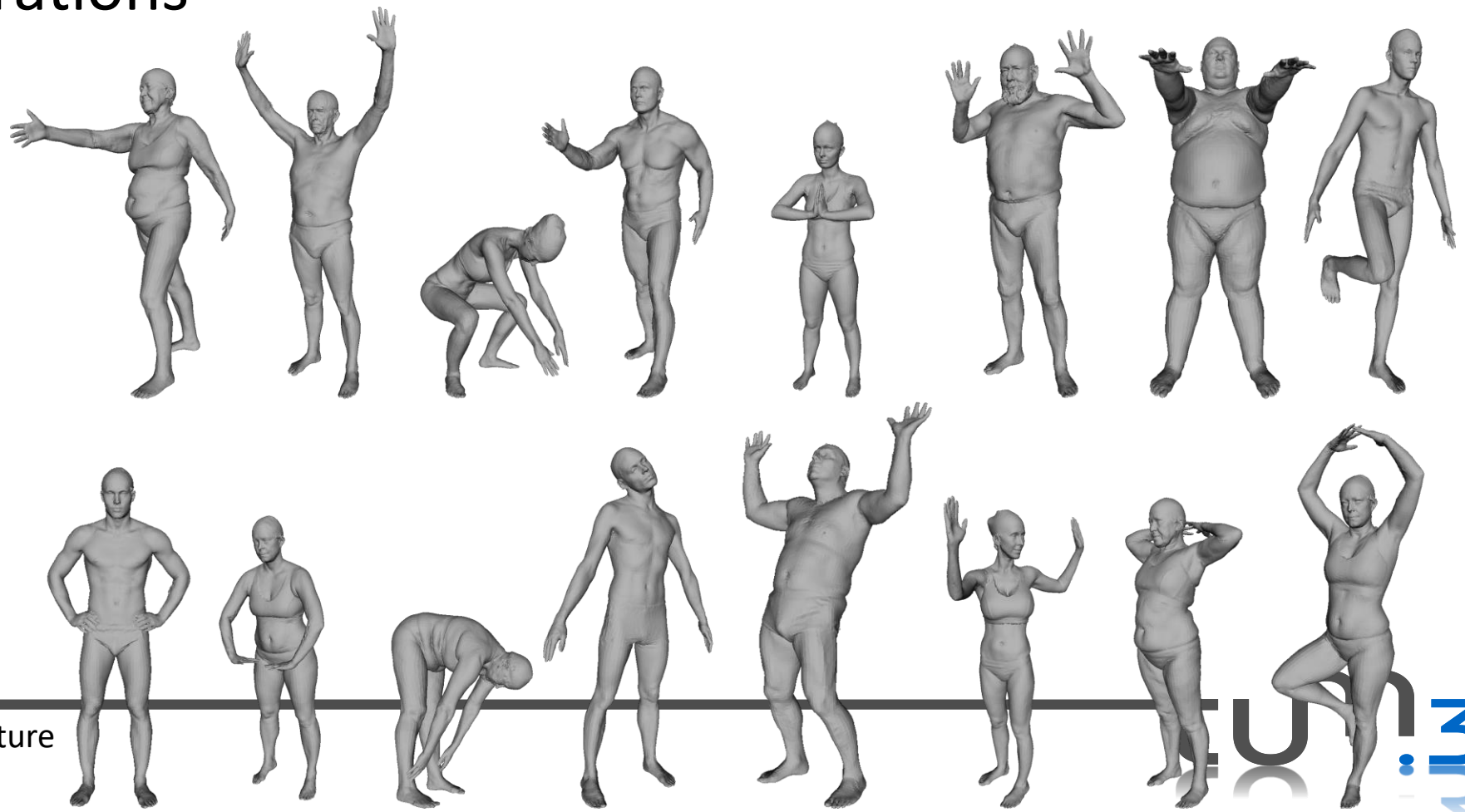
Pose Blend Shapes

- **With** blend shape correction



SMPL: Skinned Multi-Person Linear Model

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



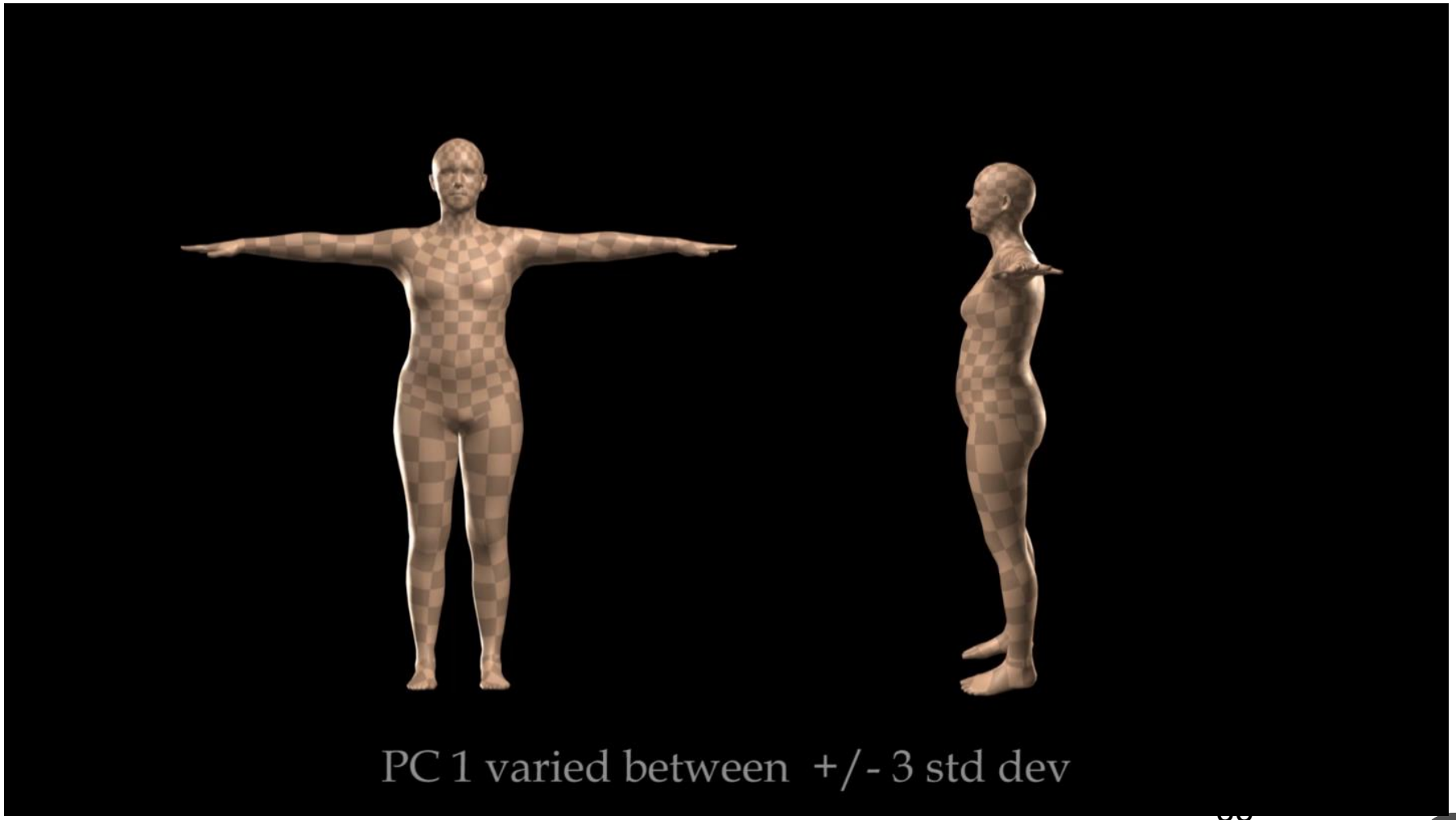
[Loper et al. 15]

SMPL: Skinned Multi-Person Linear Model

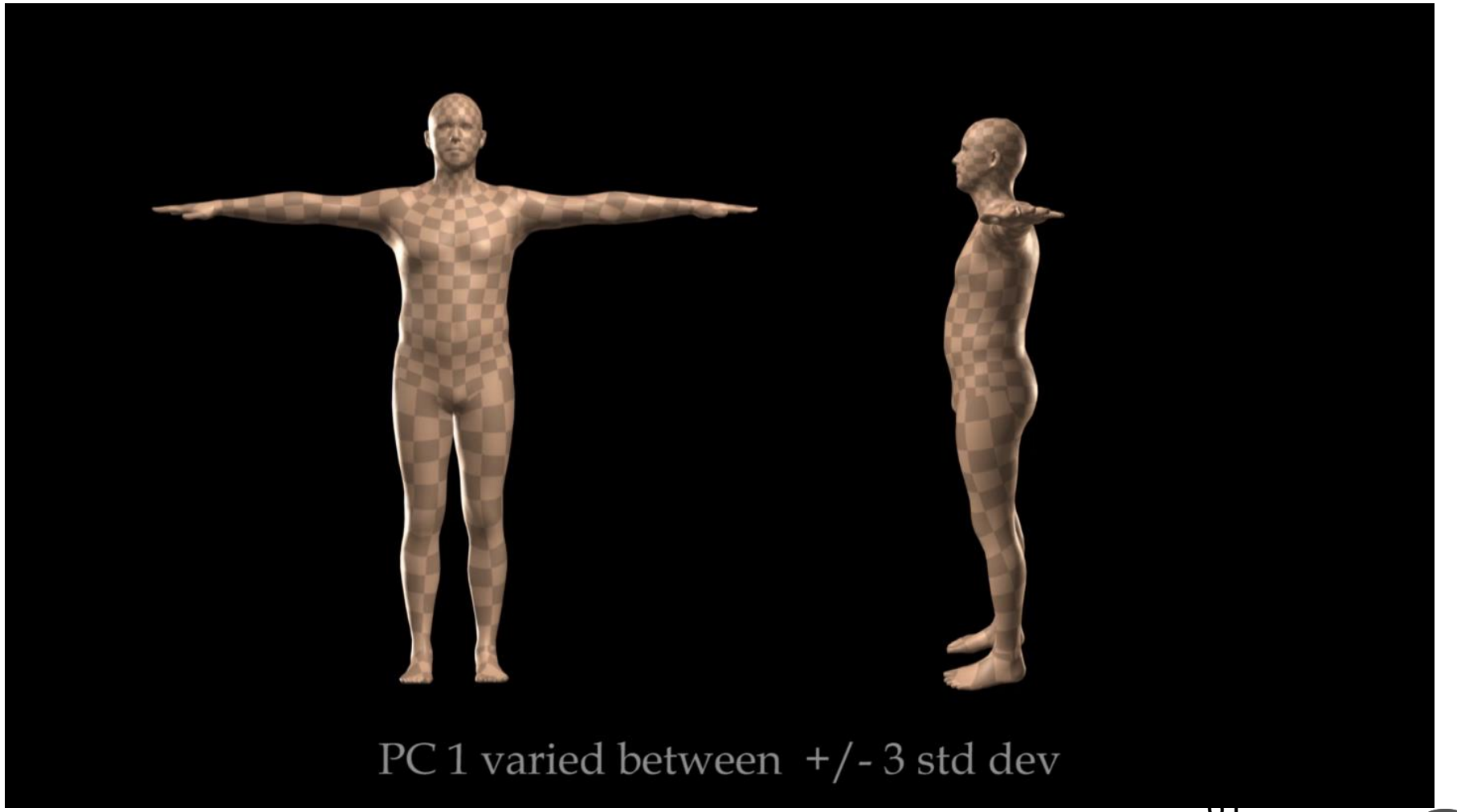
- Multishape database: PCA on ~2000 single-pose registrations per gender

[Loper et al. 15]

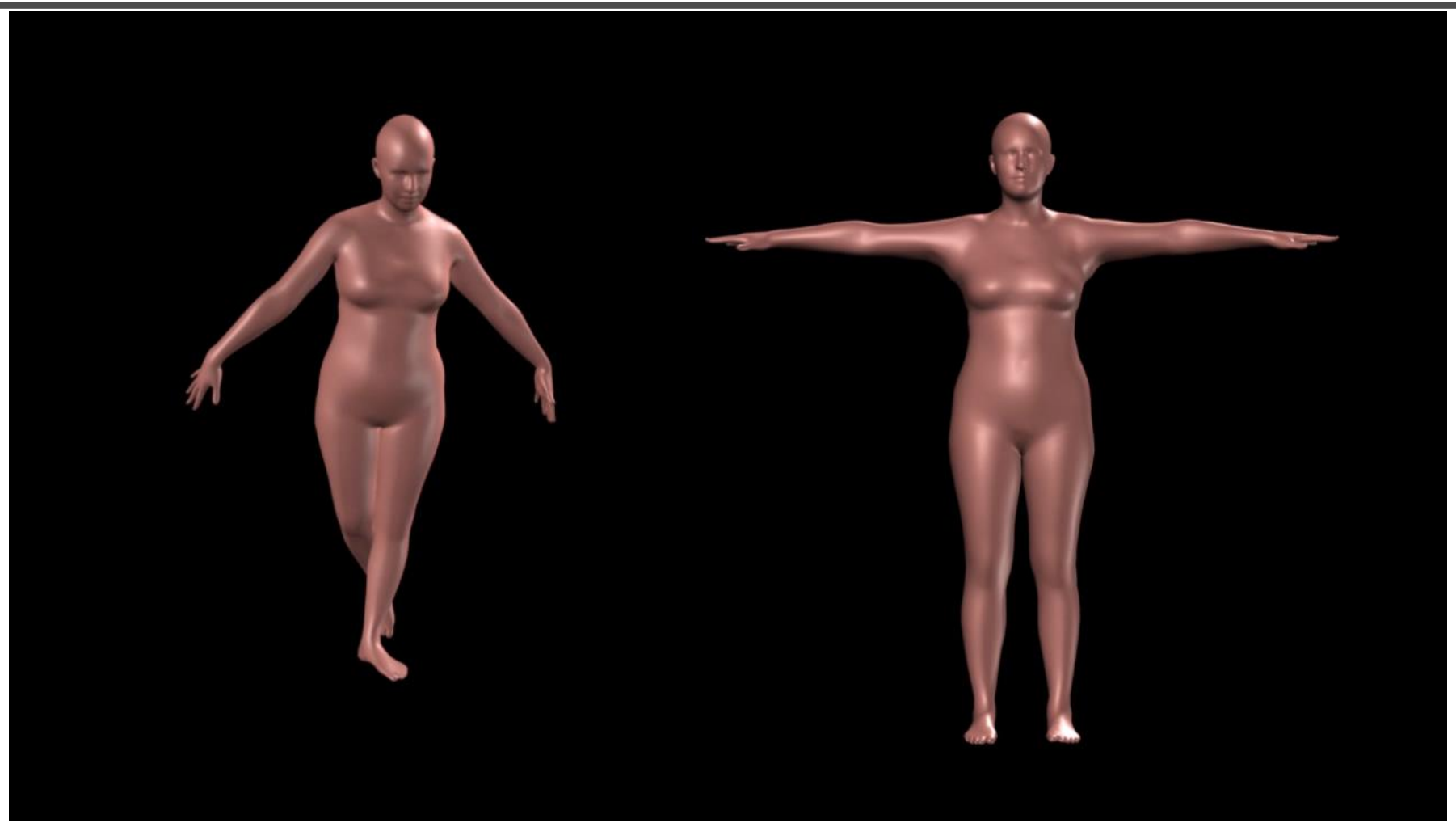
Shape Blend Shapes- Female



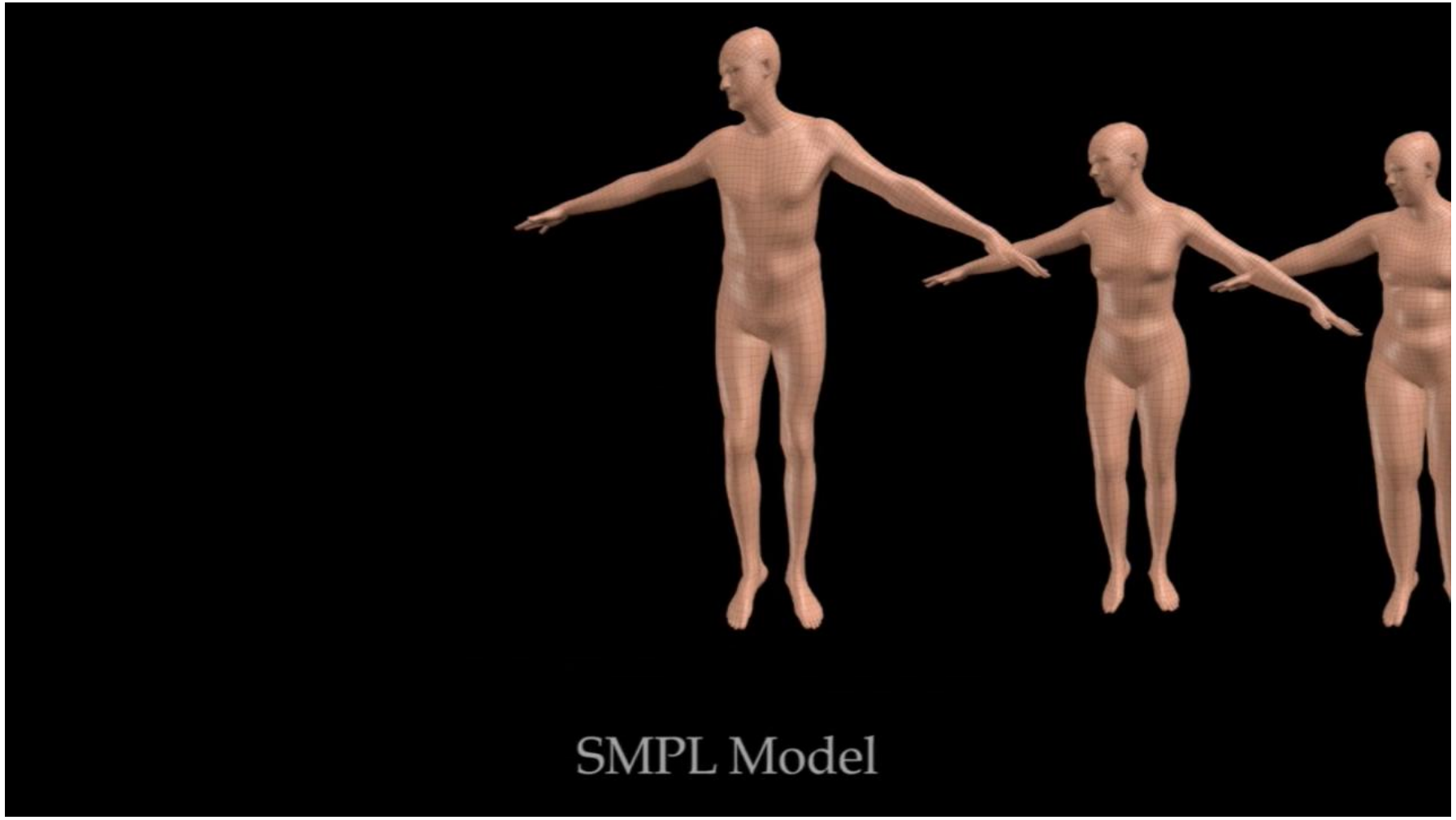
Shape Blend Shapes- Male



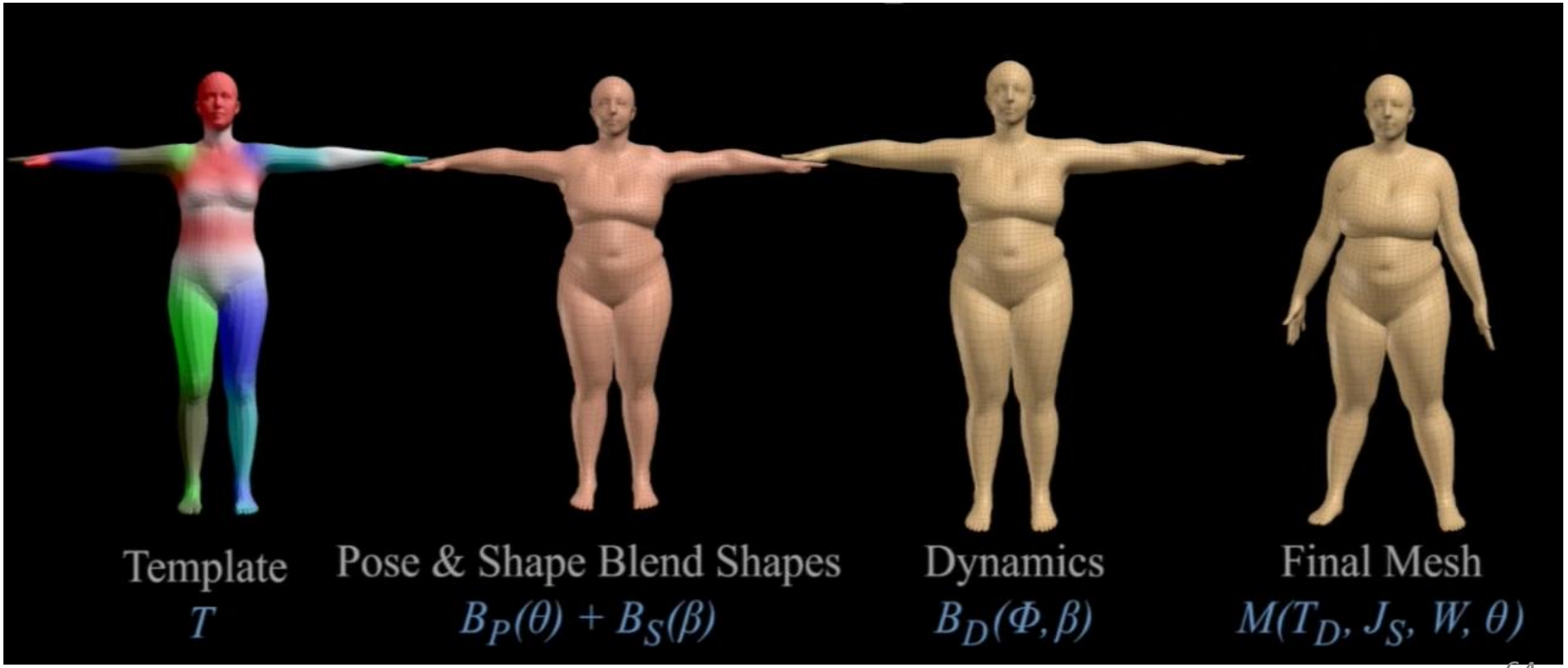
Pose Blendshapes



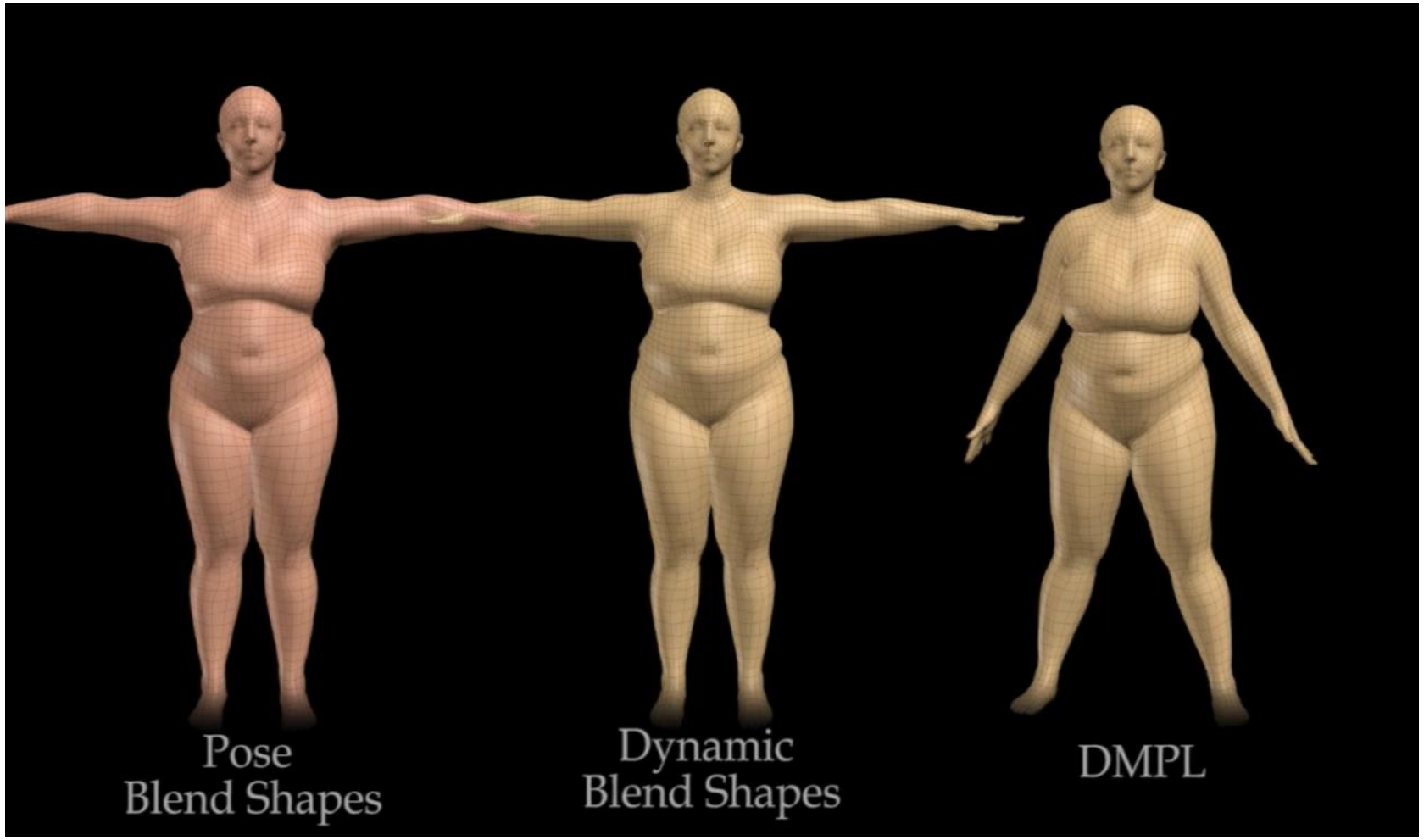
SMPL Results



Model Decomposition



Dynamics of Soft Tissue



DMPL Exaggeration

Reconstruction of Human Bodies

- Non-rigid registration to align scans
  - Fit template do 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

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

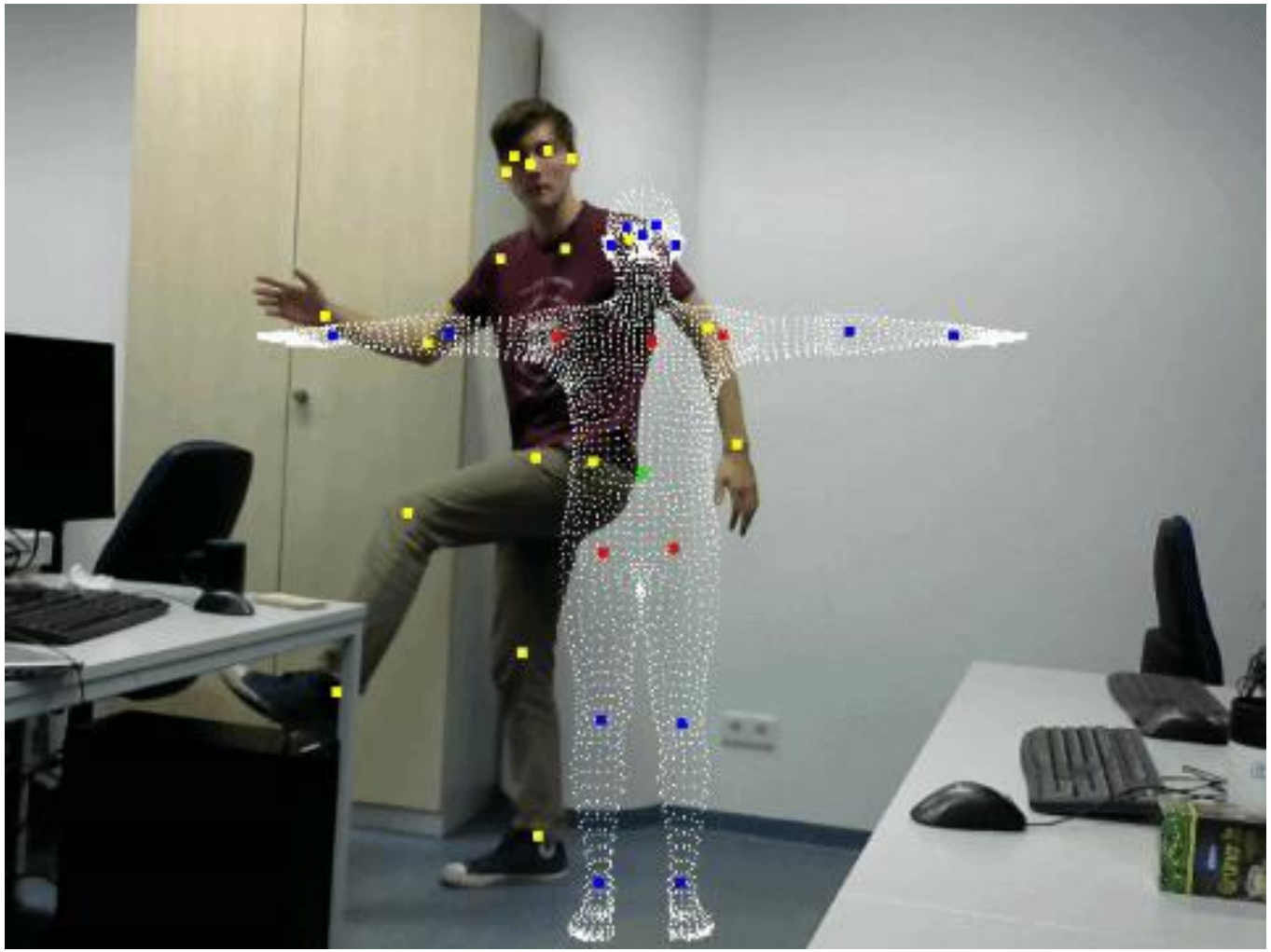
Transformation smoothness

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

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

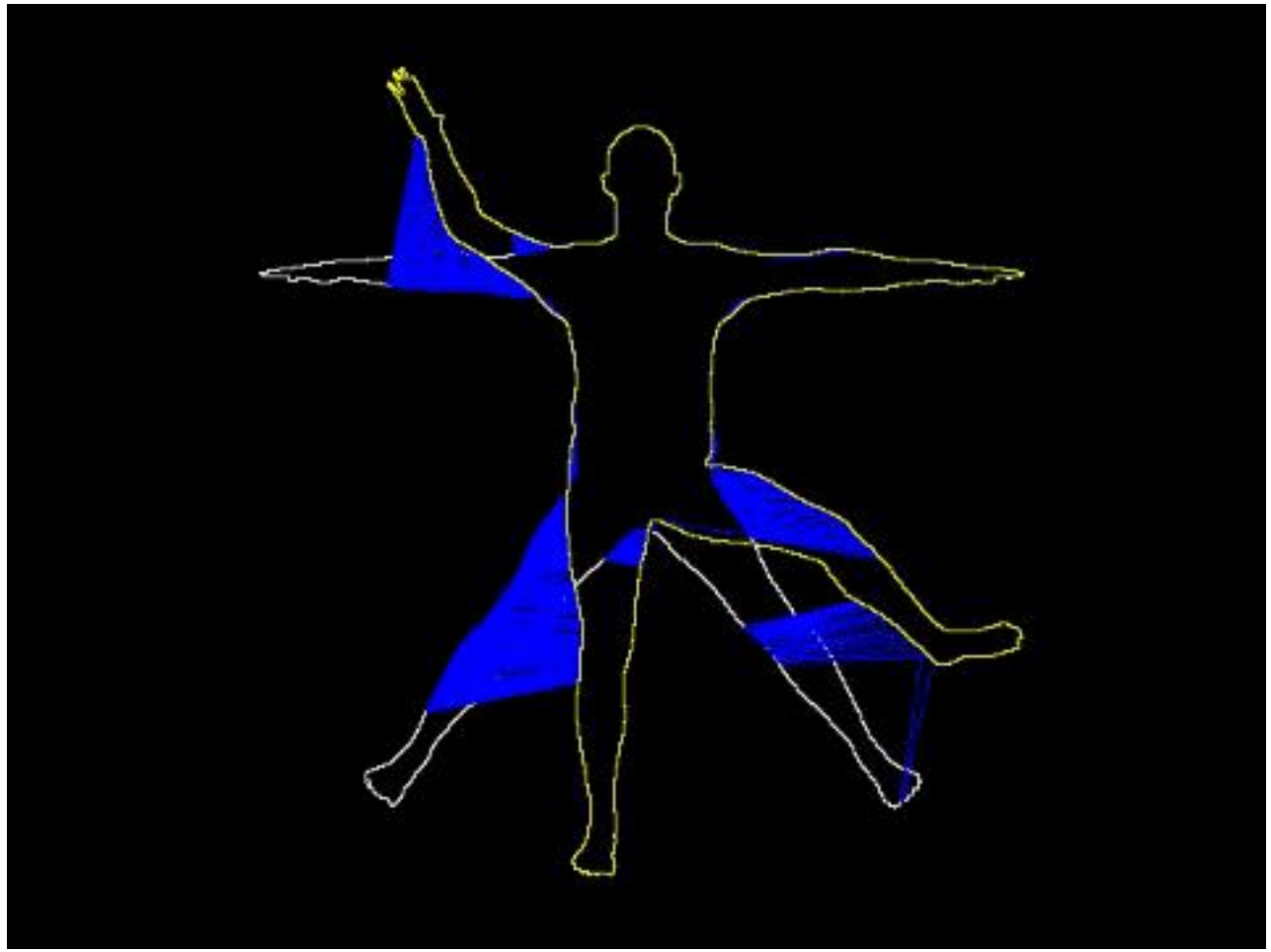
[Allen et al. 03] Slide A. Dubrovina

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

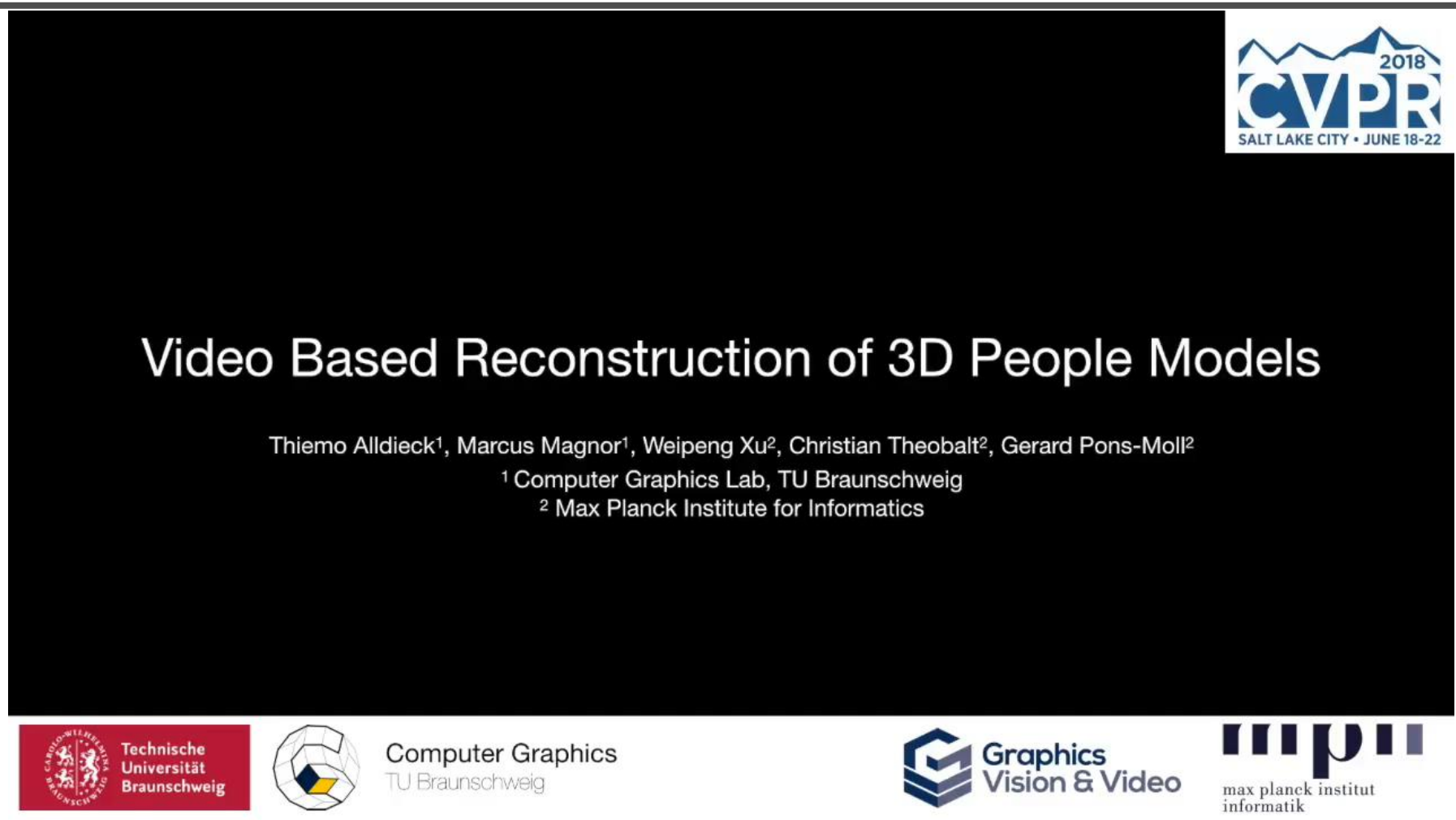


Data error: template and target surfaces as close as possible

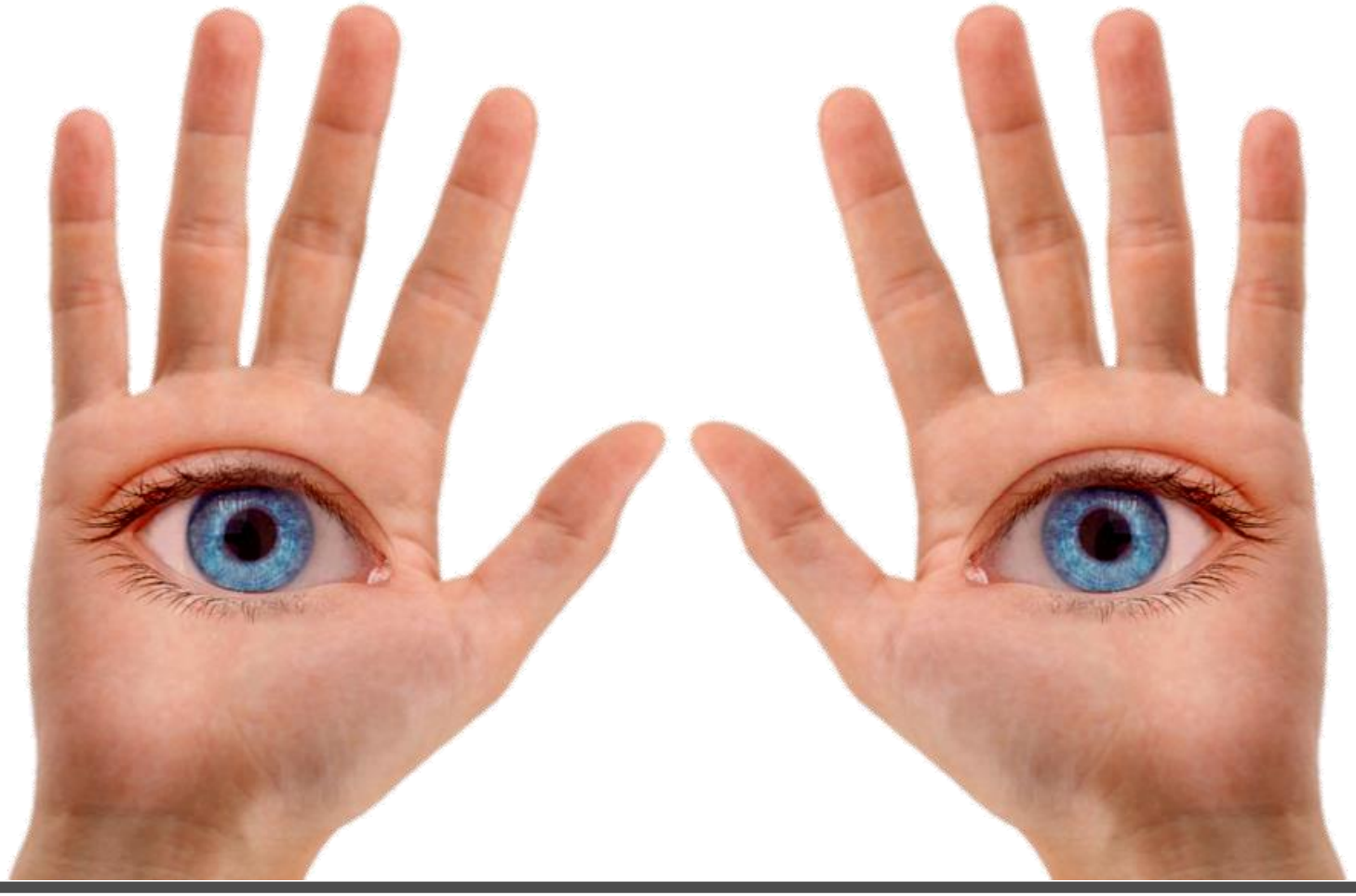
$$E_d = \sum_{i=1}^{n} w_i \, \text{dist}^2(\mathbf{T}_i \mathbf{v}_i, D)$$

Here: Silhouette constraints

Reconstruction of Human Bodies

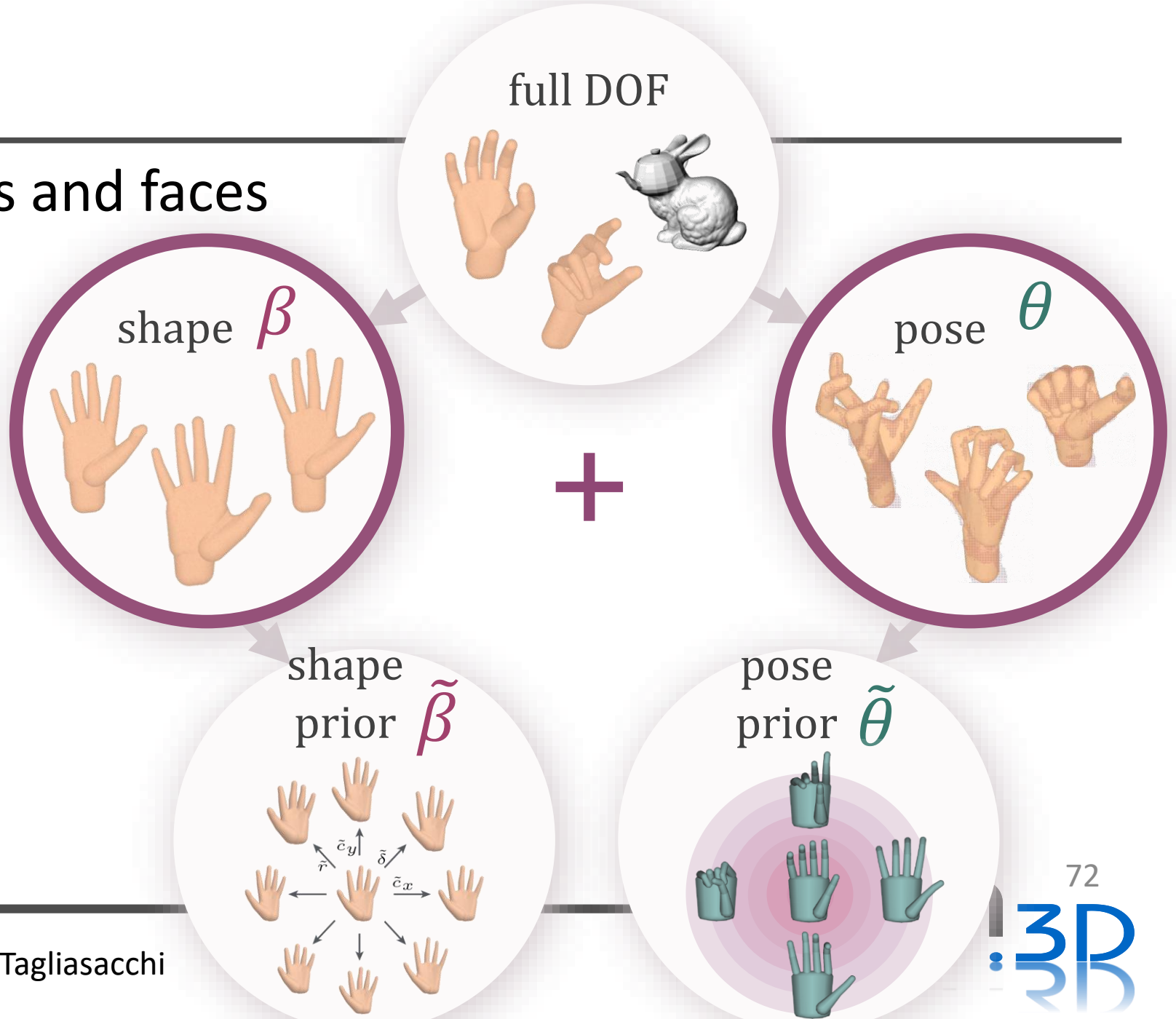


Hands

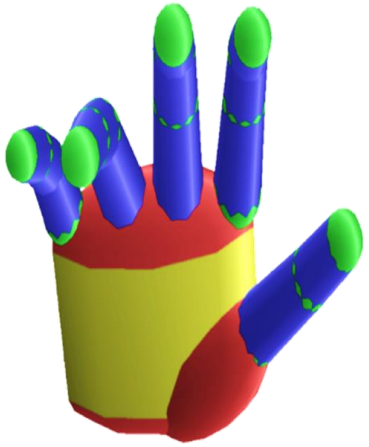


Hand Tracking

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



Hand Models

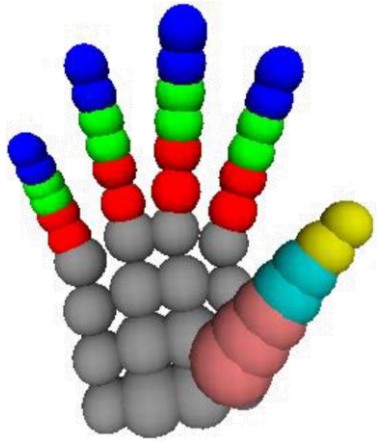


Oikonomidis et al. 2011

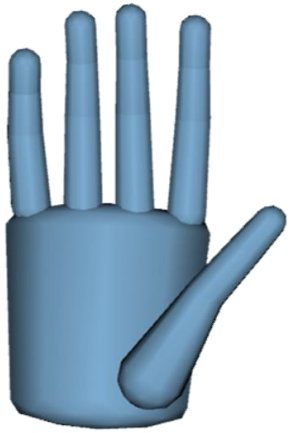
Melax et al. 2013

Sharp et al. 2015

Taylor et al. 2016

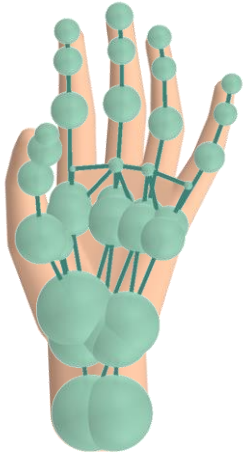


Qian et al. 2014



Tagliasacchi et al. 2015

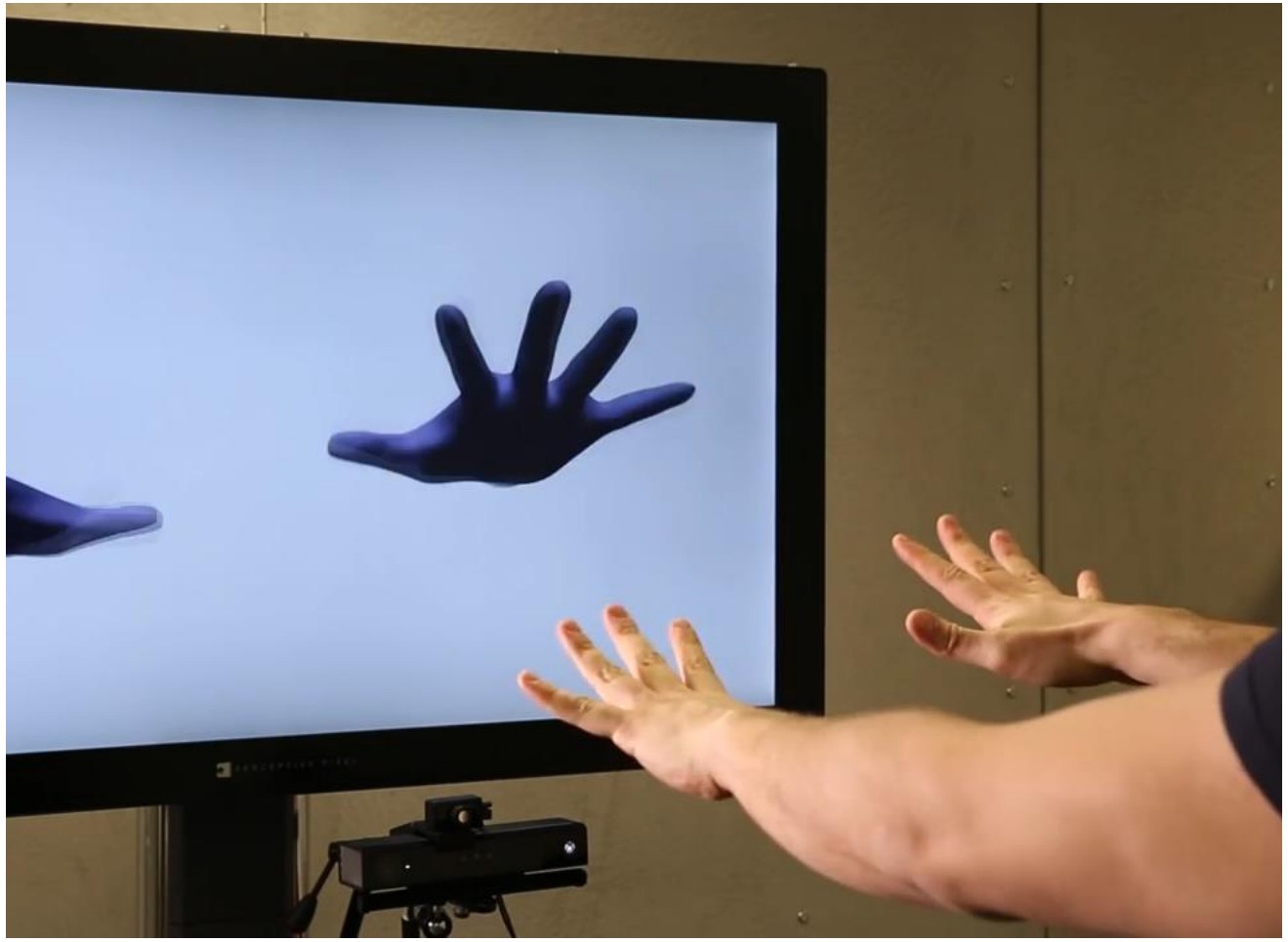
Khamis et al. 2015



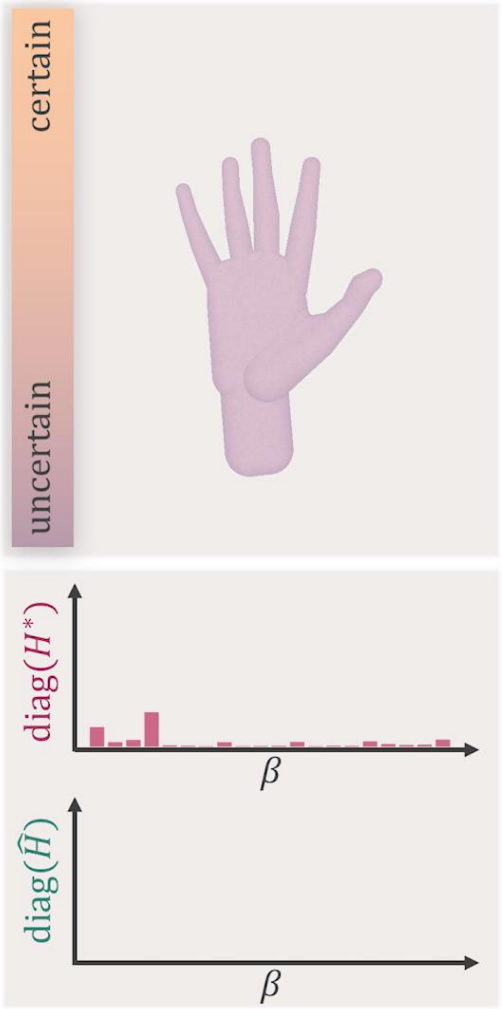
Tkach et al. 2016

[Taylor et al. 16]

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

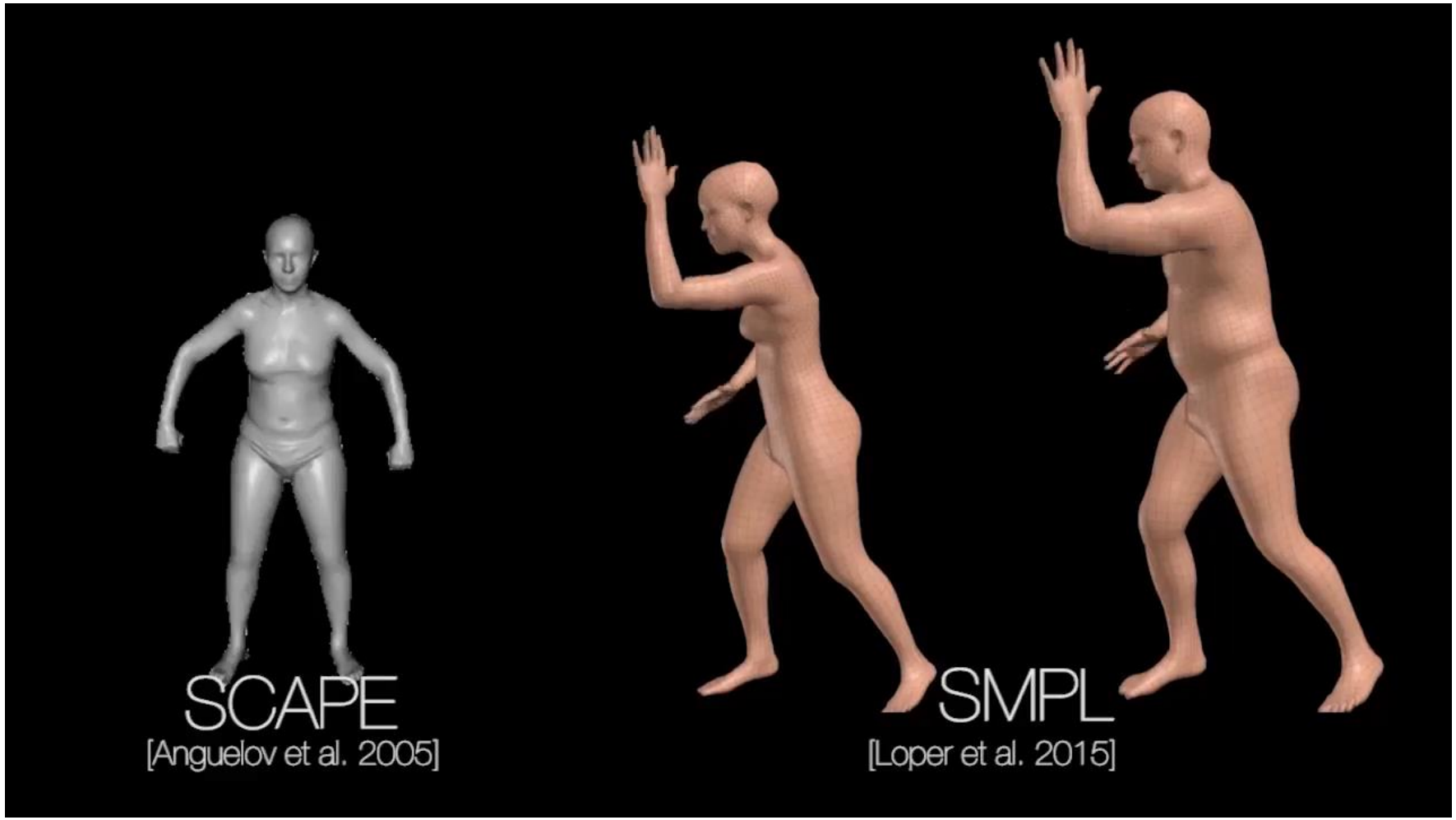


Online Calibration



[Tkach et al. 17]

Body + Hands



[Romero et al. 17]

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)

What's Left?

- Every fitting needs some form of depth data (multi-view or depth sensor)
  - I.e., not yet applicable in commodity settings

- Need init for fitting (optimization is non-convex)
  - Getting skeletons and keypoints now works reasonable well with deep learning

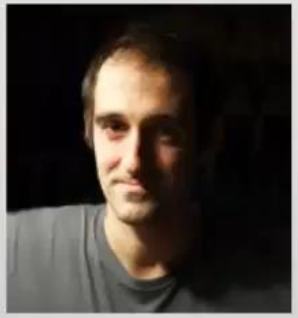
What's Left?

ClothCap:

Seamless 4D Clothing Capture and Retargeting

http://clothcap.is.tue.mpg.de/


Gerard Pons-Moll*,[1]


Sergi Pujades*,[1]


Sonny Hu[2]


Michael J. Black[1]

* Two first authors contributed equally
[1] Max Planck for Intelligent Systems
[2] BodyLabs

MAX-PLANCK-GESELLSCHAFT

BODY LABS

30 JULY – 3 AUGUST Los Angeles SIGGRAPH2017

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

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